International Journal of Neural Systems, Vol. 20, No. 2 (2010) 159–176
© World Scientific Publishing Company
DOI: 10.1142/S0129065710002334

AUTOMATIC IDENTIFICATION OF EPILEPTIC AND BACKGROUND EEG SIGNALS USING FREQUENCY DOMAIN PARAMETERS

OLIVER FAUST*, U. RAJENDRA ACHARYA† and LIM CHOO MIN‡
School of Engineering, Ngee Ann Polytechnic
Singapore 599489 Singapore
**fol2@np.edu.sg*
†aru@np.edu.sg
‡lcm@np.edu.sg

BERNHARD H. C. SPUTH
Altreonic NV, Gemeentestraat 61A Bus 1
B3210 Linden, Belgium

The analysis of electroencephalograms continues to be a problem due to our limited understanding of the signal origin. This limited understanding leads to ill-defined models, which in turn make it hard to design effective evaluation methods. Despite these shortcomings, electroencephalogram analysis is a valuable tool in the evaluation of neurological disorders and the evaluation of overall cerebral activity. We compared different model based power spectral density estimation methods and different classification methods. Specifically, we used the autoregressive moving average as well as from Yule-Walker and Burg's methods, to extract the power density spectrum from representative signal samples. Local maxima and minima were detected from these spectra. In this paper, the locations of these extrema are used as input to different classifiers. The three classifiers we used were: Gaussian mixture model, artificial neural network, and support vector machine. The classification results are documented with confusion matrices and compared with receiver operating characteristic curves. We found that Burg's method for spectrum estimation together with a support vector machine classifier yields the best classification results. This combination reaches a classification rate of 93.33%, the sensitivity is 98.33% and the specificy is 96.67%.

Keywords: Epilepsy; electroencephalogram (EEG); linear methods; spectrum estimation; Support Vector Machine (SVM); gaussian mixture model (GMM); artificial neural network (ANN).

1. Introduction

Epilepsy is a chronic neurological disorder of the brain, characterized by recurrent unprovoked seizures.[1,2] These seizures are transient signs of the disorder. The symptoms of epilepsy reach from abnormal to excessive or synchronous neuronal activity in the brain.[3] Worldwide, about 50 million people have epilepsy, with almost 90% of these people living in developing countries.[4] The disease is more likely to develop in young children or people over the age of 65 years, however, even outside this age group, it can occur at any time.[5]

Epilepsy can be diagnosed using electroencephalogram (EEG) and brain scan technology, because it affects the normal neuronal activity. Interictal, preictal and ictal are the typical stages of the epilepsy.[6] The detection of epileptic seizures from EEG data, using nonlinear methods, was proposed by Paivinen *et al.*[7] Using short sliding time windows, a set of features were computed from the data. The features came from time domain, frequency domain and nonlinear methods. They used discriminant analysis to determine the best seizure-detecting features. The outcome of their study was that the best results could be obtained by using a combination of features from both linear and nonlinear methods.

From a signal analysis perspective, EEG signals are highly complex and nonlinear in nature.

The specific signal characteristics depend on age and mental state of the subject. The symptoms of epilepsy, such as epileptic seizure, occur randomly. Therefore, the frequency of occurrence can only be estimated and stated in a statistical sense. To make an accurate forecast of eminent and future epileptic seizures implies that we understand the precise nature of the brain. This is impossible with the current state of technology. We think of the human brain as a cognitive machine which is composed out of billions of interconnected neurons. They form a network which permanently changes its state. Due to this permanent or asynchronous state change and due to the sheer complexity of the network it is impossible to understand and predict the precise state of the brain. With current technology and understanding, the best we can do is to try to estimate the state with advanced signal processing techniques and its correlation to the physiological mechanisms.

Aschenbrenner-Scheibe *et al.* proposed various methods to predict the onset of seizures based on EEG recordings.[8] Correlation dimension was used to identify preictal dimension drops up to 19 min before a seizure onset. They investigated both sensitivity and specificity, of this method, based on invasive long-term recordings from 21 patients suffering from medically intractable partial epilepsies, who underwent invasive pre-surgical monitoring. The mean length and amplitude of dimension drops showed no significant differences between interictal and preictal data sets.

Over the past two decades, much research has been done with the use of conventional temporal and frequency analysis measures in the detection of epileptic seizures from EEGs. Reasonably good results have been obtained from these studies.[9–12] Osterhage *et al.* investigate the measurements for the directionality of coupling between dynamical systems. As a case study they apply these measures to EEG signals taken from one epilepsy patient during a seizure-free interval.[13]

Adeli *et al.* have shown that the wavelet transform was particularly effective for representing various aspects of non-stationary signals.[14] They used discrete Daubechies and harmonic wavelets to analyze and characterize epileptiform discharges in patients with absence seizure. Through this decomposition, transient features were accurately captured and localized in both time and frequency domain. A combination of wavelet and chaos methodologies to detect the epilepsy using EEG analysis was studied in Refs. 15 and 16. Their results showed that, this combined approach was very effective in identifying the epilepsy.

In this paper we analyze frequency measures for the detection of epileptic activity in EEGs. The study is based on EEG data samples which are classified into three distinct classes: normal, epileptic background and epileptic seizure. We used autoregressive moving average (ARMA), Yule-Walker and Burg's method, to extract the power density spectrum (PSD) from representative EEG signal samples. Local maxima and minima were detected from these spectra. The locations of these extrema become input vectors to the classifiers. ANalysis Of VAriance between groups (ANOVA) tests on these input vectors show that the information, conveyed by these input vectors, is statistically significant. The three classifiers used here are: Gaussian mixture model (GMM), artificial neural network (ANN), and support vector machine (SVM). The different classification results are documented with confusion matrices and compared with receiver operating characteristic (ROC) curves. We found that Burg's method for spectrum estimation together with a SVM yields the best classification results. This combination reaches a classification rate of 93.33%, the sensitivity is 98.33% and the specificy is 96.67%.

Figure 1 shows the overview of the system used. Section 2 describes all the methods which were used in the individual processing blocks. Section 3 presents and discusses the results. These results are set into a wider context in Section 4. The conclusions of this paper are presented in Section 5.

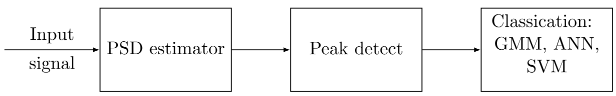


Fig. 1. Overview block diagram. The system inputs time domain EEG signals. The first processing block estimates the PSD from these signals. The local extrema are extracted with the 'Peak detect' algorithm in the next block. The location of the first 4 local maxima and the first 4 local minima forms a vector which is input to the classification block.

2. Materials and Methods

Parametric (model based) power spectrum estimation methods avoid the problem of spectral leakage and provide a better frequency resolution when compared with non-parametric methods. In general, parametric methods produce a smooth power spectral density (PSD) and the frequency bands are easily distinguishable. Furthermore, the post-processing is simpler, this makes the PSD estimation more accurate. The only drawback of parametric methods is that the model order must be chosen such that PSD estimation method yields good results for each signal class. In this study, we used three parametric PSD estimation methods, namely ARMA, Yule-Walker, and Burg. The following sections describe these methods.

Section 2.5 discusses the feature extraction. To be specific, it introduces the peak detection algorithm which states value and position of both local maxima and minima of a signal. Section 2.6 introduces ANOVA.

We use the GMM, SVM and ANN classifiers to investigate the performance of the features mentioned above in an automated pattern recognition system. Sections 2.7 – 2.9 describe these classifiers. The following section introduces the data sets which were used to obtain the results.

2.1. *Data*

The EEG data for the present study was obtained from a database available from Bonn University.[17] Gautama *et al*. discussed these datasets in Ref. 18. Three sets, (normal, epileptic background (preictal) and epileptic seizure (ictal)), of a single channel with a duration of 23.6 seconds duration, were used for the study. There are 200 data sets in both normal and preictal clases while the ictal class had 100 data sets. The normal EEG data was obtained from five healthy volunteers using a standardized electrode placement scheme, in the relaxed awake state with open eyes. In the present study, we considered only 100 data sets per class, 70 to train the classifiers and 30 to test the classifiers. The ictal EEG data was recorded during epileptic seizures from five epilepsy patients. The preictal EEG data was recorded from the same five epilepsy patients when there was no seizure. All EEG signals were recorded with the same 128 channel amplifier system, digitized with a sampling rate of 173.61 Hz and with a 12 bit A/D resolution. The data was filtered using a band pass filter with settings 0.5340 Hz $\sim$ 12 dB/octave. Figures 2 to 4 show 11.8 seconds of sample recordings for normal, preictal and ictal EEG respectively.

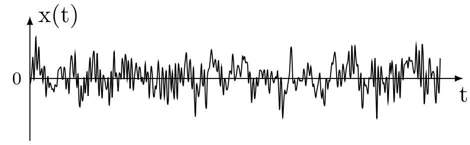


Fig. 2. Normal EEG.

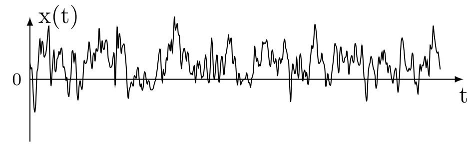


Fig. 3. Preictal EEG.

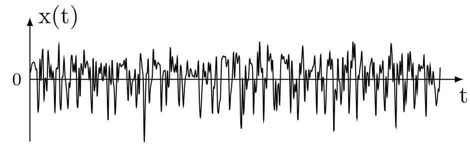


Fig. 4. Ictal EEG.

2.2. *Yule-Walker's method*

The Yule-Walker AR method of spectral estimation computes the AR parameters by forming a biased estimate of the signal's autocorrelation function and solving the least squares minimization of the forward prediction error.[19] This results in the Yule-Walker equations:

$$A \times B = C \tag{1}$$

with:

$$A = \begin{bmatrix} r_{xx}(0) & r_{xx}(-1) & \cdots & r_{xx}(-p+1) \\ r_{xx}(1) & r_{xx}(0) & \cdots & r_{xx}(-p+2) \\ \vdots & \vdots & \ddots & \vdots \\ r_{xx}(p-1) & r_{xx}(p-2) & \cdots & r_{xx}(0) \end{bmatrix}$$

and

$$B = \begin{bmatrix} \hat{a}_p(1) \\ \hat{a}_p(2) \\ \vdots \\ \hat{a}_p(p) \end{bmatrix}; \quad C = \begin{bmatrix} r_{xx}(1) \\ r_{xx}(2) \\ \vdots \\ r_{xx}(p) \end{bmatrix}$$

where r_{xx} is a biased form of the autocorrelation function. This form ensures that autocorrelation matrix, shown above, is positive definite. The biased form of the autocorrelation estimate is calculated as follows:

$$r_{xx}(m) = \frac{1}{N}\sum_{n=0}^{N-m-1} x^*(n)x(n+m) \quad m \geq 0 \tag{2}$$

The AR coefficients ($\hat{a}_p$) can be obtained by solving the of $p+1$ linear equations, extracted from Eq. 1 (for instance, by using the fast Levinson Durbin algorithm). The corresponding PSD estimate is calculated as follows:

$$\hat{P}_{\text{YW}}(f) = \frac{\sigma_{wp}^2}{|1+\sum_{k=1}^{p}\hat{a}_p(k)\, e^{-j2\pi fk}|^2} \tag{3}$$

where $\hat{\sigma}_{wp}^2$ is the estimated minimum mean square error for the pth-order, predictor calculated as follows:

$$\hat{\sigma}_{wp}^2 = \hat{E}_p^f = r_{xx}(0)\prod_{k=1}^{p}[1-|\hat{a}_p(k)|^2] \tag{4}$$

2.3. *Burg's method*

Burg's method is another algorithm to get AR model parameters. It is computationally efficient and yields a stable AR model.[20] Burg's method is based on minimizing both forward and backward prediction errors as well as estimating the reflection coefficient. The power spectrum of the pth order autoregressive process is defined as:

$$\hat{P}_{\text{Burg}}(f) = \frac{\hat{e}_p}{|1+\sum_{l=1}^{p}\hat{a}_p\, e^{-j2fl}|^2} \tag{5}$$

Where $\hat{e}_p$ denotes the total least square error. It is the sum of forward and backward prediction errors, $\hat{e}_{f,p}$ and $\hat{e}_{b,p}$ respectively. The prediction errors are calculated as follows:

$$\begin{aligned}\hat{e}_{f,p} &= x(n) + \sum_{i=1}^{p}\hat{a}_{p,i}\, x(n-i)\\ \hat{e}_{b,p} &= x(n-p) + \sum_{i=1}^{p}\hat{a}_{p,i}^*\, x(n-p+i)\end{aligned} \tag{6}$$

where $n = p+1, \ldots, U$.

One of the most important aspects to consider when using the AR method is the selection of the model order p. In this work the order of the AR model is taken as: $p = 20$.[10,21]

2.4. *ARMA method*

The ARMA model is a combination of autoregressive (AR) and moving average (MA) models.[22,23] The power spectrum of an autoregressive moving average process is given by Eq. 7.

$$\hat{P}_{\text{ARMA}}(f) = \frac{\sigma^2|\sum_{l=0}^{q}\hat{b}_p(l)e^{-j2fl}|^2}{|1+\sum_{l=1}^{p}\hat{a}_p(l)\, e^{-j2fl}|^2} \tag{7}$$

where σ^2 is the prediction error variance. Both AR coeffients ($\hat{a}_p$) and MR coeffients ($\hat{b}_p$) were obtained with the Yule Walker method as described in the previous section. In general the ARMA model is generated by filtering unit variance noise with a filter having p poles and q zeros. This method is based on the assumption that the value of the output signal depends on the previous values of the same signal (autoregressive component) and on the present and previous values of a different input signal (moving average component), plus an additional noise factor. The advantage of the ARMA model is that it can incorporate both autoregressive and moving average terms.

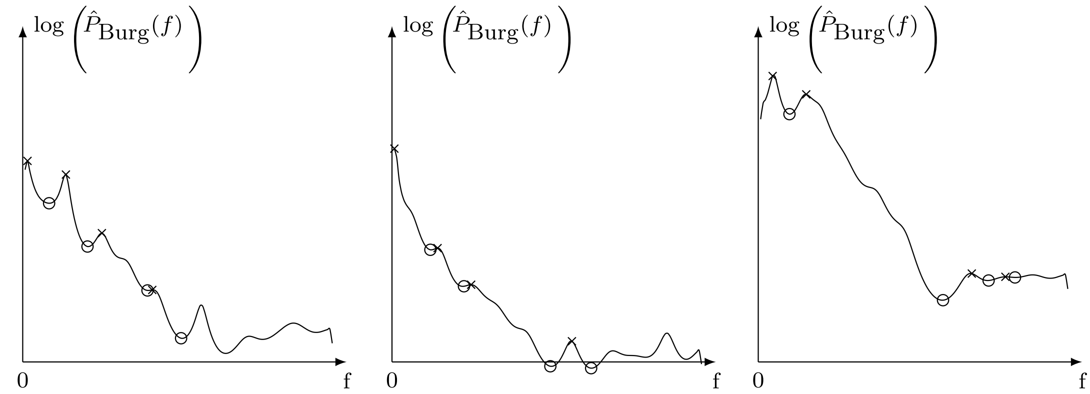
2.5. *Peak detection*

In this work we have used Billauer's[a] 'Peak detection' algorithm to locate the first 4 local maxima and the first 4 minima in the two dimensional PSD signals. Figure 5 shows the results of Burg's method of PSD estimation for normal, epileptic background and seizure signals. The local maxima are marked with a cross ($\times$), the coordinates are encoded by the amplitude $aX_{\max}$ and by the frequency $fX_{\max}$, where $X \in \{1, \ldots, 4\}$ is the number of the maxima. For example, the ordered pair $(a1_{\max}, f1_{\max})$ encodes the coordinates of the first maxima. Similarly, the local minima are marked with circles ($\circ$), the coordinates are encoded by the amplitude $aX_{\min}$ and by the frequency $fX_{\min}$, where $X \in \{1, \ldots, 4\}$ is the number of the minima.

2.6. *Analysis of variance between groups (ANOVA)*

ANOVA test uses variances to decide whether or not the *means*, which were evaluated independently for

[a] http://billauer.co.il/peakdet.html



(a) Normal (b) Epileptic Background (c) Seizur

Fig. 5. Results of local maxima and minima in the Burg PSD spectras of normal, epileptic background and seizure.

each class of input parameters, are different. The result of this test is the so called *p-value*. A low *p*-value indicates that the means of individual classes are independent. Therefore, a low *p*-value is desired for classification problems. In other words, a low *p*-value gives some certainty that it is possible to differentiate the individual classes with an automated classifier, such as GMM, ANN and SVM. Therefore, if the observed differences are high, i.e. the *p*-value is low, then the test result is considered to be statistical significant.

2.7. *Gaussian mixture model (GMM)*

GMMs have been widely used in many areas, such as pattern recognition and classification. Their use has been especially successful in speaker identification and verification.[24,25] In GMM models, a probability density function is expressed as a linear combination (with weights w_i) of N multidimensional Gaussian basis functions. Each of these basis functions is specified by its mean values μ_i and its covariance matrix Σ_i, both can be derived from the input signal. For a single observation, x, the probability density function of a given GMM model, λ:

$$p(x|\lambda) = \sum_{i=1}^{N} w_i g(s|\mu_i, \Sigma_i) \tag{8}$$

The probability density function of a single Gaussian component of D dimensions is defined as:

$$g(x|\mu_i, \Sigma_i) = \frac{1}{\sqrt{(2\pi)^D |\Sigma_i|}} \times e^{[-\frac{1}{2}(x-\mu_i)'\Sigma_i^{-1}(x-\mu_i)]} \tag{9}$$

where ($'$) denotes the vector transpose. The solution, to determine the parameters of the GMM, uses the Maximum Likelihood (ML) parameter estimation criterion. The model parameters are estimated through training, the goal is to maximize the likelihood of the observations using the so called *Expectation-Maximization* (E-M) algorithm.[26]

Usually, the initial estimates of the parameters are obtained from a sample of the training data using a simpler procedure, such as K-means.[27] The K-means procedure starts with randomly chosen initial means and assumed unit variances for the covariance matrix. This method has been adopted in this work.

2.8. *Artificial neural network (ANN)*

ANNs are comprised of densely interconnected adaptive simple processing elements called *neurons*. These neurons are interconnected, but independent entities, therefore they are capable of performing parallel computations for data processing and knowledge representation. The most commonly used neural network is called multilayer perceptron neural networks (MLPNN). We adopted MLPNNs for this study, because they operate fast and they are easy to implement. The MLPNN has been used widely for a variety of detection and estimation tasks.[28,29]

Figure 6 shows the ANN used for classification in this study. In this work, the nature of the class boundaries was not clearly known. Under these circumstances there is no theoretical method with which the network setup can be determined. By trial and error we found that a four layer network with sigmoid activation function gives good results. The input layer had 9 neurons, the two hidden layers

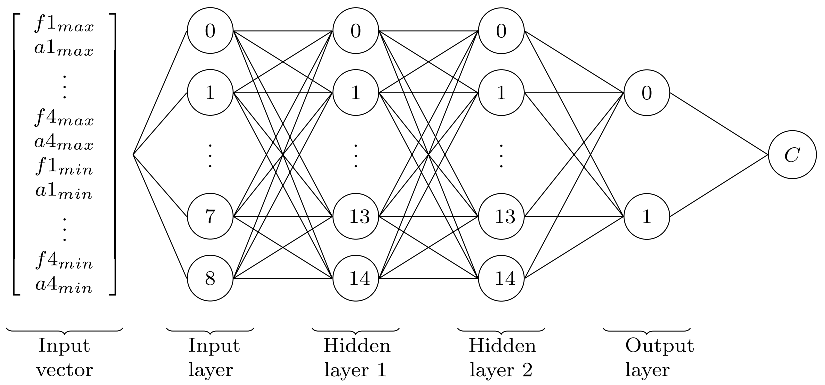


Fig. 6. MLPNN with 9 input neurons, 15 neurons in the first and second hidden layer and 2 neurons in the output layer.

have 15 neurons each and the output layer has two neurons.

The multilayer perceptron was trained with the back propagation algorithm (BPA). This is a supervised learning algorithm which aims to reduce the error between actual and desired network outputs. BPA is a so called *steepest decent method*, where weight values are adjusted in an iterative fashion while moving along the error surface to arrive at minimal range of error, when input patterns are presented to the network for learning.

During the initialization phase, the connection weights of the neural network were randomly assigned. During the training phase they are progressively modified to reduce the overall mean square error. The weight update, aimed at maximizing the rate of error reduction was set to 10^{-9}. With regards to the choice of the weight increment, there is no definite rule for its selection; however the weight increment was done in small steps. In the present case, a learning constant, $\eta = 0.9$ (that controls the step size) was chosen by trial and error.

The ideal training data set is large in size and uniformly spread throughout the class domains. In the absence of an ideal training dataset, the available data was used iteratively until the error function came down below a threshold. For quick and effective training, data was fed from all classes in a routine sequence so that the right message about the class boundaries was communicated to the ANN.

2.9. *Support vector machine (SVM)*

SVMs were initially designed for two-class problems. But, they have been extended to multi-class problems. The text below briefly explains the two-class SVM approach. The SVM operation searches for a hyperplane which acts as a decision surface that separates positive and negative values from each other with maximum margin.[30,31] This involves orienting the separating hyperplane perpendicular to the shortest line separating the convex hulls of the training data for every class, and locating it midway along this line. Let the separating hyperplane be given by $x \cdot w + b = 0$, where w is its normal. For linearly separable data $\{x_i, y_i\}$ where $x_i \in \mathbb{R}^n$ and $y_i = \{-1, 1\}, i = 1, \ldots, N$, the optimum boundary, chosen with the maximal margin criterion, is found by minimizing the objective function:

$$E = \|w\|^2 \tag{10}$$

Subject to $(x_i \cdot w + b)y_i \geq 1 \quad \forall i$.

The solution for the optimum boundary w_0 is a linear combination of a subset of the training data, $s \in \{1, \ldots, N\}$ known as the *support vectors*. This solution can be obtained more easily by translating it into its "dual form". The optimization problem can be solved by quadratic methods giving the optimum decision boundary w_0 as:

$$\mathbf{w}_0 = \sum_{<i>} \alpha_i \mathbf{y}_i \mathbf{x}_i \tag{11}$$

which is a linear combination of the support vectors with $\alpha_i \neq 0$.

Kernel functions can be used to extend the solution to nonlinear boundary problems. The dot product $(\cdot)$ in the feature space is expressed by some functions (i.e., the kernels) of two vectors in input space. The polynomial and radial basis function (RBF) kernels are commonly used. With the use of kernels, an explicit transformation of the data to the feature space is not necessary.

There are several algorithms that extend the basic binary SVM classifier to be a multi-class classifier. Examples are: the one-against-one SVM, one-against-all SVM, half against half SVM, and Directed Acyclic Graph SVM (DAGSVM). We used the RBF kernel function with a one-against-all algorithm to classify an input EEG segment among the three classes (normal, ictal and preictal). We performed an initial search for the SVM parameters by using a "grid search" approach as suggested by Hsu.[32]

2.10. *Receiver Operating Characteristic (ROC)*

The ROC curve is a plot in a two dimensional space. The x-axis is '1 - specificity' and the y-axis is 'sensitivity'. Sensitivity, also known as true positive fraction, refers to the probability that a test result is positive when a disease is present.

The area under the ROC curve indicates the classifier performance across the entire range of cut-off points. Conventionally, the area under the ROC curve must fall in the range between 0.5 and 1.[33] An area closer to 1 means that the classifier has a better accuracy. The area under the ROC curve is a good indicator for the classifier's performance.[34]

For example, Fogarty *et al.* used ROCs to analyze the tradeoff between true positive and false positive for sensor based estimates. Their case studies compare sensor-based estimates with human performance. They optimize a feature selection process for the area under the ROC curve, and they examine end-user selection of a desirable tradeoff.[35]

In this work we used ROC to test the classifiers in their ability to differentiate normal from both epileptic background and seizure. In this case, specificity measures the proportion of signals from the normal group which are correctly identified. Similarly, sensitivity measures the proportion of both epileptic background and seizure groups which are correctly identified.

3. Results

The block diagram shown Fig. 1, gives an overview on how the results were obtained. The first step is to estimate the PSD from the individual signals. We did this by applying three different methods, namely: ARMA, Yule-Walker and Burg. The local maxima and minima of the PSD curve are extracted with the peak detection algorithm. The location of the first 4 maxima and the first 4 minima forms a 16 dimensional vector which is the input for the classifiers. With respect to the local extrema location, 'first' means the extrema being located at the lowest frequency. The three classifier used in this study were: GMM, ANN and SVM. All concepts have been introduced in Section 2, therefore this section reports only the results of the individual experiments. It is structured such that there is a subsection for each PSD estimation method. Within this section, the confusion matrices of the three individual classifiers are discussed and the classifiers are compared within ROC graphs. Section 3.4 compares the results across the individual PSD estimation methods.

For all tables, presented in the following sections, C = class, N = normal, EB = epileptic background and S = seizure. 'p' stands for p-value. For all p-values, a 0 indicates a result lower than 0.0001.

3.1. *ARMA method*

Despite the fact that the ARMA method uses both autoregressive and moving average parameters, the classification results, presented in the following text, are the poorest of all three tested methods. The classification results are based on the parameters, i.e. location of the local extrema. The statistical relevance of these parameters is indicated by Table 1. It shows mean and standard deviation of both frequency f and amplitude a values for the first four maxima within each class. The last row shows the p-values from the ANOVA test. Similarly, Table 2 shows these measures for the first four local minima. The only general trend, within the ANOVA results, is that both local maxima and local minima are statistically significant.

The discussion of the classification result starts with the GMM classifier. The classification rate of the GMM classifier is 31.11%, which is below 50%. Table 3 shows the confusion matrix for this test. The numbers within the 3×3 matrix document the performance of the classifier. The first row indicates that the GMM method classifies only 3 data sets, taken from normal EEGs, correctly as normal. But, 27 normal subjects are wrongly classified as epileptic

Table 1. ARMA max: Mean and variance results calculated from individual elements of the input vector, for one class.

C	$f1_{\max}$	$a1_{\max}$	$f2_{\max}$	$a2_{\max}$	$f3_{\max}$	$a3_{\max}$	$f4_{\max}$	$a4_{\max}$
N	0.0548 ± 0.0716	0.4672 ± 0.1133	0.1738 ± 0.3074	0.0300 ± 0.1316	0.1229 ± 0.0923	0.3231 ± 0.2443	0.0374 ± 0.0945	0.0166 ± 0.1036
EB	0.2640 ± 0.1400	0.3144 ± 0.2365	0.1424 ± 0.2156	0.1230 ± 0.2362	0.2728 ± 0.1968	0.1488 ± 0.1922	0.0625 ± 0.1383	0.0364 ± 0.0824
S	0.1200 ± 0.0530	0.1200 ± 0.1756	0.0978 ± 0.2599	0.0290 ± 0.1260	0.3167 ± 0.1468	0.0858 ± 0.1342	0.0177 ± 0.0611	0.0057 ± 0.0267
p	0	0	0.1248	0.0001	0	0	0.0092	0.0197

Table 2. ARMA min: Mean and variance results calculated from individual elements of the input vector, for one class.

C	$f1_{\min}$	$a1_{\min}$	$f2_{\min}$	$a2_{\min}$	$f3_{\min}$	$a3_{\min}$	$f4_{\min}$	$a4_{\min}$
N	0.2715 ± 0.0872	0.1465 ± 0.2597	0.0282 ± 0.1209	0.0194 ± 0.0996	0.1046 ± 0.0446	0.0235 ± 0.0481	0.0117 ± 0.0513	0.0134 ± 0.0750
EB	0.2471 ± 0.1833	0.1297 ± 0.1950	0.1288 ± 0.2470	0.0791 ± 0.2022	0.1027 ± 0.1427	0.0687 ± 0.1708	0.0579 ± 0.1496	0.0368 ± 0.1251
S	0.0947 ± 0.1481	0.0956 ± 0.2591	0.0315 ± 0.1377	0.0235 ± 0.1446	0.0927 ± 0.1417	0.0194 ± 0.0725	0.0134 ± 0.0684	0.0055 ± 0.0519
p	0	0.3123	0	0.0103	0.7478	0.0025	0.0011	0.0372

Table 3. Result of GMM classification.

Target	Normal	Seizure	EB
Normal	3	27	0
Seizure	6	23	1
EB	2	26	2

seizure and no data sets were classified as epileptic background. Similarly, the second row details that 23 data sets are correctly identified as epileptic seizure. But, 6 were wrongly classified as normal and 1 was wrongly classified as epileptic background. Finally, the last row indicates that 26 data sets were correctly identified as epileptic background and 4 data sets were wrongly classified. The sum of the elements within each row is always 30, i.e. the number of data sets, within each class, used to test the classifier.

The second classifier, which was tested on the same data sets, is the ANN. This classifier achieved a classification rate of 78.89%. The confusion matrix, given in Table 4, shows that the normal data sets are better classified then with GMM. However, the classification of epileptic background is not satisfactory, because 8 out of 30 epileptic background data sets were wrongly classified as epileptic seizure.

Finally, the SVM classifier was also tested with the same data set. It achieved an even higher classification rate than the ANN classifier. To be specific, the classification rate of the SVM classifier is 85.56%. The confusion matrix, presented in Table 5, shows that SVM achieves acceptable results for all classes.

To compare the classifier performance we used ROC curves. The ROC curves, shown in Fig. 7, highlight the poor performance of the GMM classifier.

Table 4. Result of ANN classification.

Target	Normal	Seizure	EB
Normal	24	5	1
Seizure	1	25	4
EB	0	8	22

Table 5. Result of SVM classification.

Target	Normal	Seizure	EB
Normal	26	3	1
Seizure	1	25	4
EB	1	3	26

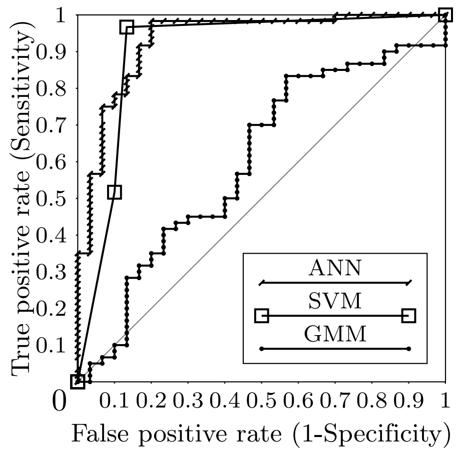


Fig. 7. ROC curves of the classifiers based on ARMA method.

Table 6 presents the statistical analysis of the ROC curves. AUC indicates the area under the curve. The area under the curve is further analyzed with the standard error (S.E.)[36] and the confidence interval (C.I.).[37] Apart from the analysis of the area under the ROC, Table 6 also provides sensitivity and specificity of the tests. These results show that the SVM classifier is the best choice for classifying the parameters obtained from the ARMA PSD.

3.2. *Yule-Walker method*

The second test sequence was conducted with the parameters extracted from the Yule-Walker PSD. Tables 7 and 8 detail the statistical significance of the parameters. The values in Table 8 support the claim that the positions of local minima are also statistically significant. Overall, the p-values are lower than the ones from the parameters obtained from the ARMA PSD.

As before, the weakest classification method is the GMM classifier. The classification rate of the GMM classifier is 81.11%. Compared with the ARMA classification rate, this is roughly 50% higher. This increase of the classification rate is also reflected in the confusion matrix, shown in Table 9. According to this table, GMM delivers poor results for normal classification, the results for the other two classes are acceptable.

Even though, the classification rate of GMM is acceptable, with a classification rate of 85.56% the ANN classifier is better. Especially the performance

Table 6. ROC analysis. For all classifier, the area under the ROC is statistically greater than 0.5.

Classifier	AUC	S.E. 95%	C.I.	Sensitivity	Specificity
GMM	0.61556	0.06114	0.49572 0.73539	0.7000	0.5333
ANN	0.93167	0.02576	0.88118 0.98215	0.9167	0.8333
SVM	0.95278	0.02110	0.91142 0.99413	0.9667	0.8667

on the normal class is more accurate, according to the confusion matrix given in Table 10.

Compared to both: GMM and ANN, the classification rate of 87.78%, achieved by the SVM classifier, is the best for parameters extracted from the Yule-Walker PSD. The SVM classifier performs good for all three classes, as shown in the confusion matrix of Table 11. This table shows that the SVM classifier yields its weakest result for seizure classification.

The ROC curves, shown in Fig. 8, describe the results of the confusion matrices. However, in this case the ROC curve does not reflect the detailed confusion matrix results. This comes from the fact that ROC curves detail only two class problems (disease present or not). According to Fig. 8 the ANN classifier is better than SVM and GMM.

Table 12 provides a detailed ROC analysis. The area under the ROC curve favors the ANN classifier, it has an area of 0.96722 compared to an area of 0.92167 for the SVM classifier. Similarly, both sensitivity and specificy of the ANN classifier are also the best, when compared to the SVM and GMM classifiers.

3.3. *Burg's method*

In general, Burg's method of spectrum estimation outperforms the other two PSD estimation techniques. Tables 13 and 14 show the statistics of these parameters. The p-values are slightly better than the ones obtained from Yule-Walker parameters. However, only the classification results provide a strong support for the claim that Burg's method is the best PSD estimation method, among the three methods, for EEG signal classification.

We achieved a classification rate of 82.22% even with the GMM classifier. The confusion matrix for the GMM classifier, provided in Fig. 15, shows a perfect classification for epileptic background. However, the classification for normal is poor and therefore the overall performance of GMM is the weakest of the three tested classifiers.

With a classification rate of 90%, the ANN classifier is better than the GMM classifier. The confusion matrix, shown in Table 16, shows that the ANN performs well for all three classes.

The best classifier for parameters obtained from Burg's PSD is the SVM with a classification rate of 93.33%. Table 17 shows the confusion matrix. The SVM performs especially well for normal classification.

Figure 9 shows the ROC curves of the three classifiers which took part in this test. The curves show that the performance of ANN and SVM is similar. They are both superior when compared to the GMM classifier.

Table 18 gives a detailed ROC analysis. The table shows that ANN has a slightly larger area under the ROC curve then SVM. However, both sensitivity and specificity are the same for ANN and SVM. Both, ANN and SVM outperform the GMM classifier in all measures.

3.4. *Comparison of the different PSD estimation methods*

This section compares the results of the SVM classifier, obtained from parameters which were extracted from different PSDs. Tables 19 and 20 summarize the ANOVA test results. These results show two trends: (1) The first extrema are statistically more significant than the following extrema, i.e. the p-value goes up towards the right side of the tables. (2) Both parameter sets, Yule-Walker and Burg, show more

Table 7. Yule-Walker max: Mean and variance results calculated from individual elements of the input vector, for one class.

C	$f1_{\max}$	$a1_{\max}$	$f2_{\max}$	$a2_{\max}$	$f3_{\max}$	$a3_{\max}$	$f4_{\max}$	$a4_{\max}$
N	0.1400 ± 0.0391	0.1346 ± 0.0209	0.2870 ± 0.1210	0.5489 ± 0.1365	0.0108 ± 0.0241	0.0210 ± 0.0272	0.0132 ± 0.0213	0.0243 ± 0.0494
EB	0.5808 ± 0.2463	0.2673 ± 0.2501	0.5986 ± 0.2787	0.6620 ± 0.3380	0.1030 ± 0.1449	0.1389 ± 0.1948	0.0373 ± 0.1185	0.0360 ± 0.1333
S	0.1758 ± 0.0767	0.3595 ± 0.2850	0.5598 ± 0.2962	0.4631 ± 0.3818	0.0098 ± 0.0259	0.0019 ± 0.0037	0.0003 ± 0.0008	0.0012 ± 0.0017
p	0	0	0.1248	0.0001	0	0	0.0092	0.0197

Table 8. Yule-Walker min: Mean and variance results calculated from individual elements of the input vector, for one class.

C	$f1_{\min}$	$a1_{\min}$	$f2_{\min}$	$a2_{\min}$	$f3_{\min}$	$a3_{\min}$	$f4_{\min}$	$a4_{\min}$
N	0.1068 ± 0.0231	0.2584 ± 0.1215	0.5099 ± 0.1266	0.6591 ± 0.1631	0.0042 ± 0.0019	0.0065 ± 0.0090	0.0196 ± 0.0554	0.0154 ± 0.0475
EB	0.2981 ± 0.2834	0.5307 ± 0.2443	0.6169 ± 0.3205	0.4467 ± 0.4348	0.1498 ± 0.2012	0.0339 ± 0.1161	0.0421 ± 0.1414	0.0785 ± 0.1724
S	0.4290 ± 0.3381	0.4885 ± 0.2451	0.4101 ± 0.3378	0.2759 ± 0.3569	0.0025 ± 0.0038	0.0002 ± 0.0004	0.0011 ± 0.0023	0.0024 ± 0.0046
p	0	0	0	0	0	0.001	0.0046	0

Table 9. Results of the GMM classification.

Target	Normal	Seizure	EB
Normal	18	8	4
Seizure	0	27	3
EB	0	2	28

Table 10. Results of ANN classification.

Target	Normal	Seizure	EB
Normal	25	4	1
Seizure	0	26	4
EB	0	4	26

Table 11. Results of SVM classification.

Target	Normal	Seizure	EB
Normal	26	1	3
Seizure	2	25	3
EB	0	2	28

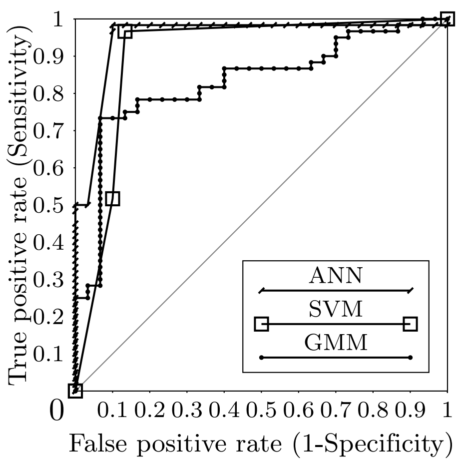


Fig. 8. The result of ROC curves of the classification based on Yule-Walker method.

statistical significance when compared to parameters obtained from the ARMA PSD.

The discussion of the ANOVA tests (p-values) gives an indication of how well the classification methods may perform. In the previous sections, both classification rate and confusion matrices show that SVM is the best classifier for the parameters obtained from the different PSDs. The ROC curves, shown in Fig. 10 indicate that Burg's method yields the best classification result. This does not contradict the ANOVA test results.

The most detailed and therefore most valuable, way of comparing the SVM performance is the combined confusion matrix. A sequence of three rows describes the SVM results of the three PSDs for the same target. The first group describes the results for normal. With 29 correctly classified data sets, obtained from the Burg PSD, the SVM classifier shows the best result. For the seizure group, the Burg PSD also yields the best SVM classification results. For epileptic background, there is tie between the SVM classification results obtained from Burg and Yule-Walker PSD.

4. Discussion

During the epilepsy state there is a sudden increase in neural discharge causing an increase in variability. These neurons, in the cerebral hemispheres, may during epilepsy mis-create abnormal electrical activity. Hence, the number of neurons available for useful information processing reduces during seizures.[38,39] So, during the seizure there is more variability, resulting in higher entropy.

Ghosh-Dastidar *et al.* have investigated automatic epilepsy and seizure detection using pattern recognition method.[40] Their proposed spiking neural network model resulted a high classifiation accuracy

Table 12. Result of ROC analysis for GMM, ANN and SVM classifiers. For all classifier, the area under the ROC is statistically greater than 0.5.

Classifier	AUC	S.E. 95%	C.I.	Sensitivity	Specificity
GMM	0.83111	0.04206	0.74867 0.91355	0.7833	0.8333
ANN	0.96722	0.01737	0.93318 1.00127	0.9833	0.9000
SVM	0.92167	0.02775	0.86728 0.97605	0.9667	0.8667

Table 13. Burg max: Mean and variance results calculated from individual elements of the input vector, for one class.

C	$f1_{\max}$	$a1_{\max}$	$f2_{\max}$	$a2_{\max}$	$f3_{\max}$	$a3_{\max}$	$f4_{\max}$	$a4_{\max}$
N	0.1400 ± 0.0391	0.1345 ± 0.0210	0.2838 ± 0.1176	0.5390 ± 0.1432	0.0112 ± 0.0260	0.0211 ± 0.0273	0.0124 ± 0.0199	0.0248 ± 0.0497
EB	0.5725 ± 0.2484	0.2623 ± 0.2448	0.5925 ± 0.2838	0.6566 ± 0.3407	0.0981 ± 0.1383	0.1434 ± 0.1971	0.0388 ± 0.1186	0.0337 ± 0.1325
S	0.1708 ± 0.0743	0.3355 ± 0.2672	0.5617 ± 0.2869	0.4247 ± 0.3661	0.0097 ± 0.0255	0.0018 ± 0.0037	0.0003 ± 0.0007	0.0011 ± 0.0015
p	0	0	0	0	0	0.0007	0.0096	0

Table 14. Burg min: Mean and variance results calculated from individual elements of the input vector, for one class.

C	$f1_{\min}$	$a1_{\min}$	$f2_{\min}$	$a2_{\min}$	$f3_{\min}$	$a3_{\min}$	$f4_{\min}$	$a4_{\min}$
N	0.1079 ± 0.0238	0.2555 ± 0.1178	0.5024 ± 0.1318	0.6427 ± 0.1764	0.0042 ± 0.0019	0.0061 ± 0.0084	0.0218 ± 0.0564	0.0139 ± 0.0462
EB	0.2961 ± 0.2814	0.5223 ± 0.2457	0.6127 ± 0.3247	0.4486 ± 0.4296	0.1522 ± 0.2022	0.0340 ± 0.1140	0.0416 ± 0.1514	0.0602 ± 0.1573
S	0.4110 ± 0.3280	0.4920 ± 0.2369	0.3717 ± 0.3160	0.3101 ± 0.3636	0.0023 ± 0.0033	0.0002 ± 0.0004	0.0011 ± 0.0022	0.0023 ± 0.0043
p	0	0.3123	0	0.0103	0.7478	0.0025	0.0011	0.0372

Table 15. Result of GMM classification.

Target	Normal	Seizure	EB
Normal	18	8	4
Seizure	0	26	4
EB	0	0	30

Table 16. Results of ANN classification.

Target	Normal	Seizure	EB
Normal	28	0	2
Seizure	0	27	3
EB	0	4	26

Table 17. Result of SVM classification.

Target	Normal	Seizure	EB
Normal	29	0	1
Seizure	0	27	3
EB	1	1	28

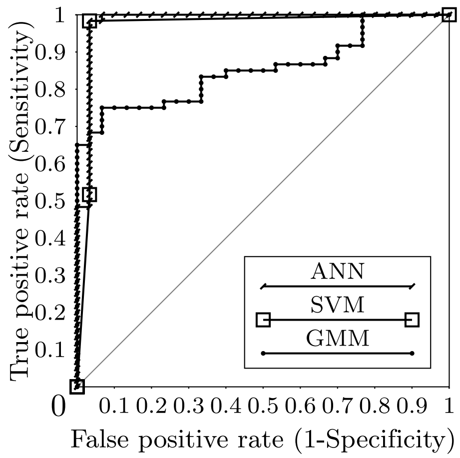


Fig. 9. The result of ROC curves of the classification based on Burg's method.

of 92.5%. In other research work, by the same leading authors, a novel principal component analysis (PCA)-enhanced cosine radial basis function neural network classifier was studied to detect the epilepsy and seixure.[41] Their method yielded a high classification accuracy (96.6%) and was robust to changes in training data with a low standard deviation of 1.4%. For epilepsy diagnosis, when only normal and interictal EEGs were considered, the classification accuracy of the proposed model was 99.3%. In their most recent work, they used Multi-Spiking Neural Network model together with a new supervised learning algorithm to identify the epilepsy and seizure.[42] The classification accuracy of this system was in the range of 90.7% to 94.8%.

Kannathal *et al.*, used different types of entropies to analyze normal and epileptic EEG signals.[43] They have successfully identified the normal and epileptic EEG signals using different entropies and neuro-fuzzy classifier with an accuracy of more than 90%.

Partial and generalized epilepsy has been detected using Radial Basis Function Neural Network (RBFNN) and Multilayer Perceptron Neural Network (MLPNNs).[44] Their studies indicate that, RBFNN (95.2%) performs better than the MLPNN (89.2%).

Recently, Chua *et al.*, have used higher order spectra (HOS) to differentiate between normal, background (preictal) and epileptic EEG signals.[45] These HOS features were fed as input to GMM and SVM classifiers for automatic identification. They have shown that, their HOS features coupled with classifiers were able to achieve 95.78% and 91.70% classification accuracy, respectively.

It was studied that, there was a significant drop in phase synchronization for the pre-ictal state.[46] In a

Table 18. ROC analysis. For all classifier, the area under the ROC is statistically greater than 0.5.

Classifier	AUC	S.E. 95%	C.I.	Sensitivity	Specificity
GMM	0.85444	0.03886	0.77828 0.93061	0.7500	0.9333
ANN	0.98222	0.01261	0.95750 1.00694	0.9833	0.9667
SVM	0.97583	0.01480	0.94683 1.00484	0.9833	0.9667

Table 19. 'p-values' of the parameters obtained from the position, i.e. value v and frequency f, of the first 4 maxima. A 0 indicates a p-value better than 0.0001.

PSD	$a1_{\max}$	$f1_{\max}$	$a2_{\max}$	$f2_{\max}$	$a3_{\max}$	$f3_{\max}$	$a4_{\max}$	$f4_{\max}$
ARMA	0	0	0.1248	0.0001	0	0	0.0092	0.0197
YW	0	0	0	0	0	0	0.0008	0.0102
Burg	0	0	0	0	0	0	0.0004	0.0149

Table 20. 'p-values' of the parameters obtained from the position, i.e. value v and frequency f, of the first 4 minima. A 0 indicates a p-value better than 0.0001.

PSD	$a1_{\min}$	$f1_{\min}$	$a2_{\min}$	$f2_{\min}$	$a3_{\min}$	$f3_{\min}$	$a4_{\min}$	$f4_{\min}$
ARMA	0	0.3123	0	0.0103	0.7478	0.0025	0.0011	0.0372
YW	0	0	0	0	0	0.001	0.0046	0
Burg	0	0	0	0	0	0.0007	0.0096	0

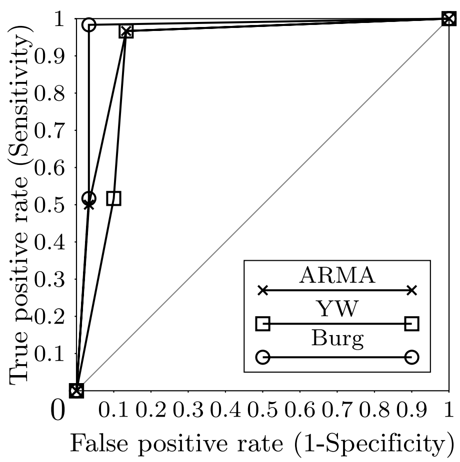


Fig. 10. Performance of the SVM classifiers.

Table 21. Area under ROC curves.

Classifier	ARMA	YW	Burg
GMM	29	0	1
SVM	0	27	3
ANN	1	1	28

Table 22. Comparison of the SVM confusion matrices from the three different spectrum estimation methods. In all cases, Burg's method outperforms the other two PSD estimation methods.

Target	PSD method	Result		
		Normal	Seizure	EB
Normal	ARMA	26	3	1
	YW	26	1	3
	Burg	29	0	1
Seizure	ARMA	1	25	4
	YW	2	25	3
	Burg	0	27	3
EB	ARMA	1	3	26
	YW	0	2	28
	Burg	1	1	28

controlled study, they were able to predict seizures up to an accuracy of 70% cases with no false positives in the control groups.

The chaotic features like Correlation Dimension, Hurst exponent, Lyapunov exponant and approximate entropy can be used to characterize the signal. These features extracted were used for automatic diagnosis of seizure onsets which would help the patients to take appropriate precautions.[47] These nonlinear features were used to train both Gaussian mixture model (GMM) and support vector machine (SVM) classifiers. Their results show that the GMM classifier performed better with average classification efficiency of 95%, sensitivity and specificity of 92.22% and 100% respectively.

It can be seen from our results that, AR Burg's method coupled with SVM performs better than the other combinations. It is able to identify the unknown class with a specificity is 98.33% and the sensitivity is 96.67%, which is comparable with other nonlinear methods.

During the last decade, electrical stimulation has been used to treat several neurologic disorders such as epilepsy.[48] Recently, therapeutic stimulation of epileptic foci has attracted interest in the research community.[49] We hope that accurate classification of normal, preictal and ictal mental states will improve integrated epilepsy treatment system, such as the ones proposed by Shoeb *et al.* and Osorio/Frei.[50,51]

5. Conclusion

EEG signals can be used to discriminate and subsequently diagnose different brain states, like normal, epileptic background and epileptic seizure. Changes in the EEG signals might be quite prominent, as in the case of an epileptic seizure or more hidden (complex), as in the case of epileptic background. In the time domain, only a trained eye can detect the different states. This work shows that, characteristics of these different mental states are also visible in the spectral domain. We used three different parametric PSD estimation methods (ARMA, Yule-Walker and Burg's) to estimate the power distribution in the frequency domain. The 'Peak detection' algorithm was used to extract local maxima and minima. The location of these extremas formed a vector which was input to the classifiers. The performance of the classifiers was stated with confusion matrices and they were compared with ROC curves.

The comparison of both the different PSD estimation methods and the different classification methods showed that the combination of Burg's method and SVM classifier yields the best results, with a specificity of 98.33% and a sensitivity of 96.67%.

This signal classification is another step towards an automated system that is able to diagnose different mental conditions based on EEG signals. Such a system would significantly improve clinical workflows, because it frees up trained personal from routine jobs.

References

1. Commission on Epidemiology and Prognosis, International League Against Epilepsy, Guidelines for epidemiologic studies on epilepsy, Guidelines for epidemiologic studies on epilepsy. Commission on Epidemiology and Prognosis, International League Against Epilepsy, *Epilepsia* **34**(4) (1993) 592–596.
2. W. Blume, H. Lüders, E. Mizrahi, C. Tassinari, B. W. van Emde and J. Engel, Glossary of descriptive terminology for ictal semiology: Report of the ILAE task force on classification and terminology, *Epilepsia* **42**(9) (2001) 1212–1218.
3. R. Fisher, B. W. van Emde, W. Blume, C. Elger, P. Genton, P. Lee and J. Engel, Epileptic seizures and epilepsy: Definitions proposed by the International League Against Epilepsy (ILAE) and the International Bureau for Epilepsy (IBE), *Epilepsia* **46**(4) (2005) 470–472.
4. World Health Organization, Epilepsy: Aetiogy [sic], epidemiology and prognosis, Published online, last accessed 1 Sept 2009 (2001). http://www.who.int/mediacentre/factsheets/fs165/en/
5. The National Society for Epilepsy, What is Epilepsy, Published online, last accessed 14 June 2007 (2001). http://www.epilepsynse.org.uk/AboutEpilepsy/Whatisepilepsy.
6. H. Adeli and S. Ghosh-Dastidar, (in cooperation with N. Dadmehr), *Automated EEG-based Diagnosis of Neurological Disorders — Inventing the Future of Neurology* (CRC Press, Taylor & Francis, Boca Raton, Florida, 2010).
7. N. Paivinen, S. Lammi, A. Pitkanen, J. Nissinen, M. Penttonen and T. Gronfors, Epileptic seizure detection: A nonlinear viewpoint, *Computer Methods and Programs in Biomedicine* **79**(2) (2005) 151–159.
8. 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* **126**(12) (2003) 2616–2626.
9. R. S. Fisher, W. R. Webber, R. P. Lesser, S. Arroyo and S. Uematsu, High-frequency EEG activity at the start of seizures., *J Clin Neurophysiol* **9**(3) (1992) 441–8.
10. M. Akin and M. K. Kiymik, Application of periodogram and AR spectral analysis to EEG signals, *J. Med. Syst.* **24**(4) (2000) 247–256.
11. J. Gotman, *Handbook of Clinical Neurophysiology* (*Vol.* 3, *chapter* 2.11) *Automatic detection of epileptic seizures, Presurgical Assessment of the Epilepsies with Clinical Neurophysiology and Functional Imaging*, Elsevier, 2004, pp. 155–165.
12. H. Lee, A. Cichocki and S. Choi, Nonnegative Matrix Factorization for Motor Imagery EEG Classification, *International Journal of Neural Systems* **17**(4) (2007) 305–317.
13. H. Osterhage, F. Mormann, T. Wagner and K. Lehnertz, Measuring the directionality of coupling: Phase versus state space dynamics and application to EEG time series, *International Journal of Neural Systems* **17**(3) (2007) 139–148.
14. H. Adeli, Z. Zhou and N. Dadmehr, Analysis of EEG records in an epileptic patient using wavelet transform, *Journal of Neuroscience Methods* **123**(19) (2003) 69–87.
15. H. Adeli, S. Ghosh-Dastidar and N. Dadmehr, A wavelet-chaos methodology for analysis of EEGs

and EEG sub-bands to detect seizure and epilepsy, *IEEE Transactions on Biomedical Engineering* **54**(2) (2007) 205–211.
16. 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* **54**(9) (2007) 1545–1551.
17. EEG time series Database, URL (last accessed 09.09.2009): http://www.meb.unibonn.de/epileptologie/science/physik/eegdata.html.
18. T. Gautama, D. P. Mandic and M. M. Van Hulle, Indications of nonlinear structures in brain electrical activity, *Phys. Rev. E* **67**(4) (2003) 046204.
19. P. Stoica and R. Moses, *Introduction to spectral analysis* (Prentice-Hall of India Pvt. Ltd, Upper Saddle River, NJ, 1997).
20. E. D. Übeyl and İnan Güler, Spectral analysis of internal carotid arterial Doppler signals using FFT, AR, MA, and ARMA methods, Computers in biology and medicine **34**(4) (2004) 293–306.
21. İnan Güler, M. Kiymik, M. Akin and A. Alkan, AR spectral analysis of EEG signals by using maximum likelihood estimation, *Computers in biology and medicine* **31**(6) (2001) 441–450.
22. S. Mukhopadhyay and P. Sircar, Parametric modelling of non-stationary signals: A unified approach, *Signal Process* **60**(2) (1997) 135–152.
23. Task Force of the European Society of Cardiology the North American Society of Pacing Electrophysiology, Heart Rate Variability: Standards of Measurement, Physiological Interpretation, and Clinical Use, *Circulation* **93**(5) (1996) 1043–1065. http://circ.ahajournals.org.
24. D. Reynolds, T. Quatieri and R. Dunn, Speaker verification using adapted Gaussian mixture models, *Digital Signal Processing* **10**(23) (2000) 19–41.
25. C. Seo, K. Y. Lee and J. Lee, GMM based on local PCA for speaker identification, *Electronics Letters* **37**(24) (2001) 1486–1488.
26. J. Bilmes, *A Gentle Tutorial on the EM Algorithm and its Application to Parameter Estimation for Gaussian Mixture and Hidden Markov Models* (1997).
27. T. Kanungo, D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman and A. Y. Wu, An efficient k-means clustering algorithm: Analysis and implementation, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **24**(7) (2002) 881–892.
28. A. T. Tzallas, M. G. Tsipouras and D. I. Fotiadis, Automatic seizure detection based on time-frequency analysis and artificial neural networks, *Intell. Neuroscience* (2007) 1–13.
29. F. Takens, Detecting strange attractors in turbulence, *Dynamical Systems and Turbulence* (1981), 366–381.
30. V. N. Vapnik, The Nature of Statistical Learning Theory (Springer-Verlag New York, Inc., New York, NY, USA, 1995).
31. V. Vapnik, *Estimation of Dependences Based on Empirical Data: Springer Series in Statistics* (Springer Series in Statistics) (Springer-Verlag New York, Inc., Secaucus, NJ, USA, 1982).
32. C. W. Hsu, C. C. Chang and C. J. Lin, A practical guide to support vector classification, Tech. rep., National Taiwan University, Taipei (2003). URL http://www.csie.ntu.edu.tw/cjlin/papers/guide/guide.pdf.
33. J. DeLeo, Receiver Operating Characteristic Laboratory (ROCLAB): Software for developing decision strategies that account for uncertainty management in artificial neural network decision-making, in *Proceedings of Second International Symposium on Uncertainty Modeling and Analysis* (1993), pp. 141–144.
34. T. Downey, D. Meyer, R. Price and E. Spitznagel, Using the receiver operating characteristic to assess the performance of neural classifiers. Neural Networks, *Computers in Biology and Medicine* **5** (1999) 3642–3646.
35. J. Fogarty, R. S. Baker and S. E. Hudson, Case studies in the use of roc curve analysis for sensor-based estimates in human computer interaction, in GI '05: *Proceedings of Graphics Interface 2005, Canadian Human-Computer Communications Society*, School of Computer Science, University of Waterloo, Waterloo, Ontario, Canada (2005), pp. 129–136.
36. M. Akin and M. K. Kiymik, A simple approximation for unbiased estimation of the standard deviation, *The American Statistician* **25**(4) (1971) 30–32.
37. H. Goldstein and M. J. R. Healy, The graphical presentation of a collection of means, *Journal of the Royal Statistical Society. Series A* (Statistics in Society) **158**(1) (1995) 175–177.
38. J. W. Sleigh, E. Olofsen, A. Dahan, J. Goede de and A. Steyn-Ross, Entropies of the EEG: The effects of general anaesthesia, in *Proceedings of the 5th International Conference on Memory, Awareness and Consciousness* (2001), in press.
39. U. R. Acharya, O. Faust, N. Kannathal, T. J. Chua and S. Laxminarayan, Dynamical analysis of EEG signals at various sleep stages, *Computer Methods and Programs in Biomedicine* **80**(1) (2005) 37–45.
40. S. Ghosh-Dastidar and H. Adeli, Improved spiking neural networks for EEG classification and epilepsy and seizure detection, *Integr. Comput.-Aided Eng.* **14**(3) (2007) 187–212.
41. S. Ghosh-Dastidar, H. Adeli and N. Dadmehr, Principal component analysis-enhanced cosine radial basis function neural network for robust epilepsy and seizure detection, *IEEE Transactions on Biomedical Engineering* **55**(2) (2008) 512–518.

42. S. Ghosh-Dastidar and H. Adeli, A new supervised learning algorithm for multiple spiking neural networks with application in epilepsy and seizure detection, *Neural Networks* **22** (2009), in press.
43. N. Kannathal, U. R. Acharya, C. Lim and P. Sadasivan, Characterization of EEG – A comparative study, *Computer Methods and Programs in Biomedicine* **80**(1) (2005) 17–23.
44. K. Aslan, H. Bozdemir, C. Şahin, S. N. Oğulata and R. Erol, A radial basis function neural network model for classification of epilepsy using EEG signals, *J. Med. Syst.* **32**(5) (2008) 403–408. doi: http://dx.doi.org/10.1007/s10916-008-9145-9.
45. K. C. Chua, V. Chandran, U. R. Acharya and C. M. Lim, Automatic identification of epileptic EEG signals using higher order spectra, *International Journal of Engineering in Medicine* **223**(4) (2009) 485–495.
46. F. Mormann, T. Kreuz, C. Rieke, R. G. Andrzejak, A. Kraskov, P. David, C. E. Elger and K. Lehnertz, On the predictability of epileptic seizures., *Clinical Neurophysiology* [official journal of the *International Federation of Clinical Neurophysiology*] **116**(3) (2005) 569–587.
47. U. R. Acharya, K. C. Chua, T. C. Lim, D. Tay and J. S. Suri, Autoamtic identification of epileptic eeg signals using nonlinear parameters, *Journal of Mechanics in Medicine and Biology* in press.
48. C. Hamani, D. Andrade, M. Hodaie, R. Wennberg and A. Lozano, Deep Brain stimulation for the treatment of epilepsy, *International Journal of Neural Systems* **19**(3) (2009) 213–226.
49. A. L. Velasco, F. Velasco, M. Velasco, J. M. Nunez, D. Trejo and I. García, Neuromodulation of epileptic foci in patients with non-lesional refractory motor epilepsy, *International Journal of Neural Systems* **19**(3) (2009) 139–147.
50. A. Shoeb, J. Guttag, T. Pang and S. Schachter, Non-invasive computerized system for automatically initiating vagus nerve stimulation following patient-specific detection of seizures or epileptiform discharges, *International Journal of Neural Systems* **19**(3) (2009) 157–172.
51. I. Osorio and M. Frei, Seizure abatement with single DC pulses: Is phase resetting at play?, *International Journal of Neural Systems* **19**(3) (2009) 149–156.