CLASSIFICATION OF TRANSIENT EVENTS IN EEG RECORDINGS USING SUPPORT VECTOR MACHINES

A.T. Tzallas*, C. D. Katsis*, P. S. Karvelis**, D. I. Fotiadis**, S. Konitsiotis*** and S. Giannopoulos***

*Dept. of Medical Physics, Medical School, University of Ioannina, GR 451 10 Ioannina, Greece
**Unit of Medical Technology and Intelligent Information Systems, Dept. of Computer Science, University of Ioannina, GR 451 10 Ioannina, Greece
***Dept. of Neurology, Medical School, University of Ioannina, GR 451 10 Ioannina, Greece

me00716@cc.uoi.gr, me00526@cc.uoi.gr, pkarvel@cs.uoi.gr, fotiadis@cs.uoi.gr, skonitso@cc.uoi.gr, sgiannop@cc.uoi.gr

Abstract: The detection of epileptic spikes (ESs) occurring in the electroencephalogram (EEG) between seizures is vital in the diagnosis of epilepsy. However, automated epileptic spike detection methods based on this approach suffer from false detections due to the presence of numerous types of artefacts (muscle activity, eye blinking activity) and non-epileptic waveforms (sharp alpha activity) usually referred as transient events. In this paper, we introduce an automated method which detects transient events in EEG recordings and classifies those as epileptic spikes (ESs), muscle activity (EMG), eye blinking activity (EOG) and sharp alpha activity (SAA). The proposed methodology involves: (i) signal preprocessing and transient event detection, (ii) clustering of transient events, (iii) feature extraction and (iv) classification of transient events using support vector machines (SVMs). Other classification schemes were tested as well. Our methodology was evaluated on data from 25 subjects and the best obtained overall accuracy is 84.83%.

Keywords: **EEG, Epilepsy, Clustering, Spike Detection, Support Vector Machines**

Introduction

Electroencephalography (EEG) provides a direct measure of cortical activity with millisecond temporal resolution. EEG is the recording of the electrical potentials generated by the cerebral cortex nerve cells. There are two types of EEG depending on where the signal is taken in the head: scalp or intracranial. For scalp EEG, which is the focus of this study, small metal discs, known as electrodes, are placed on the scalp with good mechanical and electrical contact. Intracranial EEG is obtained by special electrodes implanted in the brain during the surgery. The EEG has been found to be a valuable tool in the diagnosis of numerous brain disorders. Nowadays, the EEG recording is a routine clinical procedure and is particularly useful in the investigation of epilepsy [1].

Epilepsy is a disorder of brain function that affects about 1% of the population. It is characterized by sudden recurrent and transient disturbances of mental function, caused by excessive discharge of groups of brain cells. In the case of epilepsy two categories of abnormal activity are recognized in the EEG recordings: the inter-ictal activity which takes place between seizures (when the patient does not have seizures) and the ictal activity which takes place during an epileptic seizure. The most common forms of the inter-ictal EEG activity are: the individual or isolated spikes, the sharp wave and the spike wave complex [1-3]. Throughout this paper, no distinction is made among spike, sharp wave and spike wave complexes and therefore they are collectively termed epileptic spikes (ESs).

The detection of epilepsy can be achieved by visual scanning of inter-ictal EEG recordings, for ESs by an experienced EEGer. However, visual review of the vast amount of EEG data has serious drawbacks. Visual inspection is prohibitively very time consuming and inefficient, especially in the case of long recordings. In addition disagreement among the EEGers on the same recording is possible due to the subjective nature of the analysis. Thus computer-assisted analysis becomes quite necessary in practice [4-7].

To date, many automated detection algorithms have been developed. They are based on: (i) mimetic [2,8-9], (ii) template matching [10], (iii) parametric [11], (iv) artificial neural network [3,5,6,12-14] and (v) knowledge–based rules approaches [5,6,12,13,15-17]. Those methods recognize features to detect ESs using objective criteria. Each method has some unique advantages, but none of them alone can fulfil the requirement of ES detection. This is due to the fact ESs are similar to waves which are part of the background activity (SAA) and to artefacts (EMG and EOG). The majority of the reported works address only ES detection and only a few have been previously applied to classify transients events [16,18,19].

In this paper we describe an automated method based on a four-stage schema (Fig.1), which detects transient events in EEG recordings and classifies those as ES, EMG, EOG and SAA. In the first stage, a data driven segmentation algorithm is used to eliminate areas of low background activity and candidate transient

 ISSN: 1727-1983

events are detected by a windowing procedure. In the second stage, transient event clusters are automatically identified and for each transient event cluster its prototype shape is accurately determined. In the third stage, sixteen time-domain and frequency-domain features of each prototype transient event are extracted. Finally in the fourth stage, the prototype transient events are classified as epileptic spike (ES), muscle activity (EMG), eye blinking (EOG) and sharp alpha activity (SAA) by means of a Support Vector Machine (SVM). Two other classification schemes, Linear Discriminant Analysis (LDA) and Quadratic Discriminant Analysis (QDA) were also tested but reported lower accuracy. Our approach is novel since it does not require a priori knowledge of the number and the shape of transient event clusters. Furthermore, it is fully automatic and no additional operations are required by the neurologist.

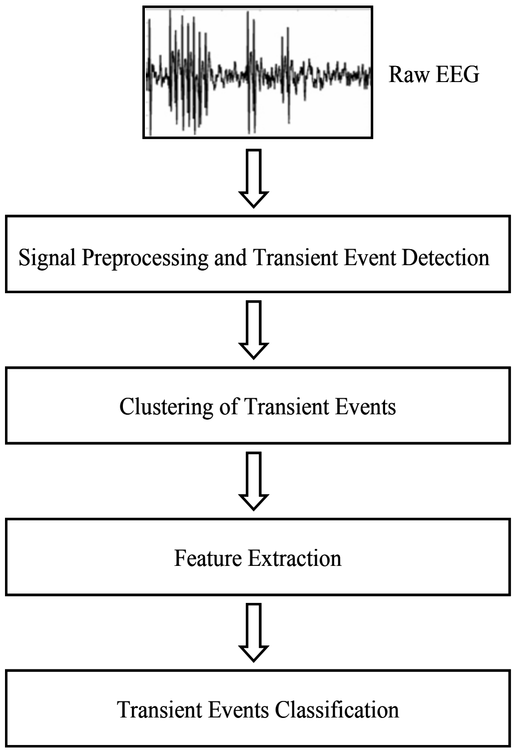

Figure 1: The proposed four-stage methodology

Materials and Methods

A. Signal Preprocessing and Transient Event Detection

Our approach is based on the fact that a transient event appears as a peak in the EEG recording. Therefore, the first stage is the elimination of low background activity and the extraction of the peaks from the EEG recording. To eliminate areas of low background activity, we first choose a threshold T which depends on the mean absolute value of the the whole EEG signal in each single channel each montage. The threshold T is calculated as follows:

$$T = \frac{1}{N} \cdot \sum_{i}^{N} |x_i| \qquad (1)$$

where x_i represents the discrete input values and N is the number of samples.

This threshold is used to identify peaks in the EEG signal. Windows with a constant length of 91 samplng points (355 msec) is used, centred at each identified peak†. If a larger peak‡ is found in the window, the window is centred at this peak; otherwise the window of 91 signal points is considered as candidate transient event. The procedure is illustrated in Fig.2.

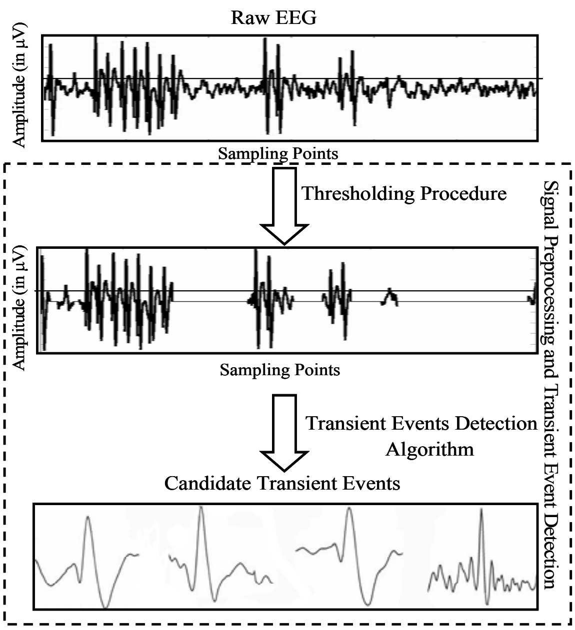

Figure 2: A schematic representation of the first stage

B. Clustering of Transient Events

The clustering of the transient events performs as a pre-classifier (assigning each transient event to a cluster) and thus, not only reduces the computation time but also increases the overall detection performance [20]. In this stage, transient event clusters are automatically detected and for each transient event cluster its prototype shape is accurately determined. The detection of the number of clusters in EEG recordings is based on the minimization of the regularized cost function *C(x,y)* with respect to the distance of the candidate transient events from the cluster centres (first term in Eq.(2)) and with respect to the distance of the cluster centres from each other (second term in Eq. (2)):

$$C(x,y) = \sum_{i=1}^{p} \sum_{j=1}^{k} f(y^{(j)} \mid x^{(i)}) \, \| x^{(i)} - y^{(j)} \|^2 + \sum_{i=1}^{p} \sum_{j=1}^{k} \tilde{I}_j \, \tilde{f}(y^{(j)} \mid x^{(i)}) \, \| y^{(j)} - y^{(\omega)} \|^2, \qquad (2)$$

where $\tilde{I}_j$ is a multiplier unique to each cluster and

† The length of the window has been selected using expert knowledge about the size of each transient event e.g. a spike has duration 20-70 msec, a sharp wave lasts 70-200 msec, an EMG wave is less than 30 msec and an EOG wave lasts more than 150 msec.

‡ The largest peak in a window is called vertex.

$$f(y^{(j)} | \; x^{(i)}) = \begin{cases} 1 & if \quad j = \arg\min_m \left\| x^{(i)} - y^{(m)} \right\|^2, \\ 0 & otherwise \end{cases} \quad (3)$$

$$\tilde{f}(y^{(j)} | \; x^{(i)}) = \begin{cases} 1 & if \quad y^{(i)} \in N_{y^{(\omega)}}, \; \omega = \operatorname{argmin}_m \left\| x^{(i)} - y^{(m)} \right\|^2 \\ 0 & otherwise \end{cases} \quad (4)$$

are indicator functions. $N_{y^{(\omega)}}$ is the neighbourhood of the cluster center $y^{(\omega)}$, $x(i) \in \Re^n$ is a pattern - in our case a transient event - p is the number of patterns $\{x(i): i = 1,2,\ldots,p\}$ and k the number of clusters and $\arg\min_m ||x^{(i)}-y^{(m)}||^2$ returns the argument which minimizes the norm $||x^{(i)}-y^{(m)}||^2$. More details about the method can be found in [21].

Using k-means clustering we start choosing an initial set of prototypes. This implies the partition of the patterns into k clusters. Each cluster is represented by a prototype, which is computed as the centroid of the patterns belonging to that cluster. The steps of the clustering procedure are shown below.

Step 1		Set initial parameters:
	1.1	k, initial number of clusters
	1.2	T_p, minimum number of patterns is one cluster,
	1.3	T_d, minimum distance between two centroids (inter-cluster distance)
Step 2		Partition the space of the patterns into k clusters minimizing the cost function $C(x, y)$.
Step 3		Delete clusters having $p < T_p$.
Step 4		Combine clusters having $d < T_d$.
Step 5		Iteration of steps 2-4, until no significant change of the number of prototypes is reached.

C. Feature Extraction

To our knowledge, time-domain and frequency-domain features of the EEG waveforms have been used in the literature [3,6,7,12-19]. In this work, we use sixteen features of each prototype transient event, which are defined as (Fig. 3):

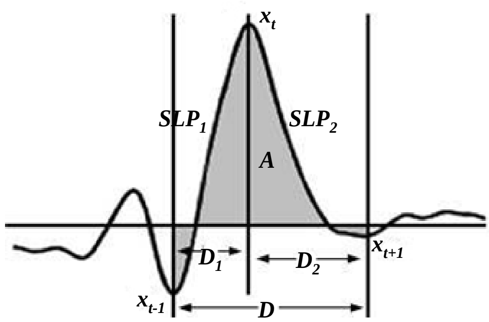

Figure 3: Features extracted from each transient event

1) Duration (D): D_1 and D_2 represent the duration of each transient event before and after the vertex x_t. D represents the sum of D_1 and D_2. The D_1 and D_2 durations are measured from the vertex to the point where the slope changes rapidly (turning point). This means that the duration is measured at the point where there is more than a 60% drop in the slope or a change in the direction of the slope.

2) Area (A): It is the area below the curve for the calculated duration (grey region in Fig. 3).

3) Average slope ($ASLP$): It is given as:

$$ASLP = \frac{1}{2}(|SLP_1| + |SLP_2|), \quad (5)$$

where $SLP_1 = x_t - x_{t-1}$ and $SLP_2 = x_{t+1} - x_t$ are the slopes of the lines connecting the vertex and the two turning points (Fig. 3).

4) Sharpness (SH): It is the changing rate of the slope at the vertex point. If the vertex point is denoted as x_t, SH can be calculated as:

$$SH = |(x_{t+1} - x_t) - (x_t - x_{t-1})| \Rightarrow$$
$$SH = |SLP_2 - SLP_1|. \quad (6)$$

5) Standard Deviation (STD): It is defined as:

$$STD = \sqrt{\frac{1}{L}\sum_{t=1}^{L}(y_{\mathrm{i}} - \mu_{\mathrm{x}})^2}, \quad (7)$$

where y_i represents the discrete input values, L is the number of samples for each prototype transient event (L=91) and $\mu_y = \sum_{i=1}^{L} \frac{y_i}{L}$ is the mean.

6) Dominant frequency (DF): In order to estimate the Power Spectrum Density (PSD) of the transient event, an autoregressive model (AR) of order 10 is used. The frequency where the maximum amplitude of the PSD was observed is the dominant frequency (DF).

7-16) Discrimination of Power Spectrum Density ($DPSD$): The PSD of each transient event is divided into ten distinct frequency ranges. The average of the PSD for each one of these is called Discrete Power Spectrum Density (DPSD).

D. Classification of Transient Events

The classification of transient events into predefined classes (ES, EMG, EOG and SAA) is achieved using Support Vector Machines (SVMs) [22,23], Linear Discriminant Analysis (LDA) [24] and Quadratic Discriminant Analysis (QDA) [25].

D_1. Support Vector Machine (SVM): It is considered as a state-of-the-art classifier for both linear and non-linear classification. SVMs belong to the family of kernel based classifiers. SVMs implicitly map the data into the feature space where a hyperplane (decision boundary) separating the classes may exist. This implicit mapping is achieved with the use of kernels, which are functions that return the scalar product in the feature space by performing calculations in the data space. The simplest case is a linear SVM trained to classify linearly separable data. After re-normalisation,

ISSN: 1727-1983 © 2005 IFMBE

the training data, $\{x_i, y_i\}$ for i=1, … , m and $y_i \in \{-1,1\}$, must satisfy the constraints in Eq. (8) and (9), where w is a vector containing the hyperplane parameters and b is an offset.

$$x_i w + b \geq +1 \text{ for } y_i = +1, \tag{8}$$

$$x_i w + b \leq -1 \text{ for } y_i = -1. \tag{9}$$

The points, for which the equalities in the above equations are satisfied and have the smallest distance to the decision boundary, are called support vectors. The distance between the two parallel hyperplanes on which the support vectors for the respective classes lie is called the 'margin'. Thus, the SVM finds a decision boundary that maximises the margin (Fig. 4).

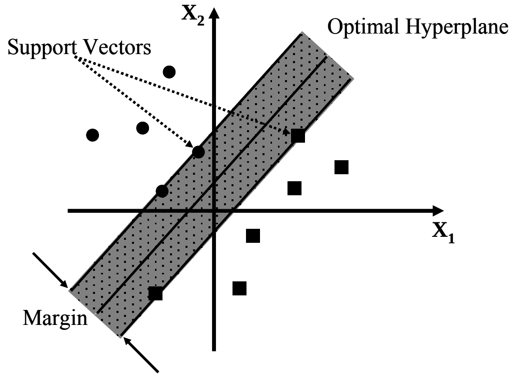

Figure 4: A hyperplane that maximizes the separating margin between two classes (indicated by data points marked by "■"s and "●"s). Support vectors lie on the boundary hyperplanes of the two classes.

Finding the decision boundary, then it becomes a constrained optimization problem which must minimize $||w||^2$ subject to the constraints (Eq. (8) and (9)) and is solved using Lagrange multipliers. The general solution is given by Eq. (10).

$$f(x) = \sum_i a_i y_i \langle x_i, x \rangle, \tag{10}$$

where α_i are Lagrange multipliers.

In the case of non-linear classification, kernels are used to map the data into a higher dimensional feature space in which linear classification may be possible. The general solution will then be of the form shown in Eq. (11). Depending on the choice of the kernel function, SVMs can provide both linear and non-linear classification.

$$f(x) = \sum_i a_i y_i \mathrm{K}\langle x_i, x \rangle. \tag{11}$$

Many implementations of kernels can be found in literature, whereby four popular are:

- Linear:

$$\mathrm{K}(x_i, x) = x_i^T x \ . \tag{12}$$

- Polynomial:

$$\mathrm{K}(x_i, x) = (\gamma x_i^T x \ + r)^d, \gamma>0. \tag{13}$$

- Radial Basis Function (RBF):

$$\mathrm{K}(x_i, x) = \exp(-\gamma \left\| x_i - x \right\|^2), \gamma>0. \tag{14}$$

- Sigmoid:

$$\mathrm{K}(x_i, x) = \tanh(\gamma x_i^T x \ + r). \tag{15}$$

Where γ, r and d are kernel parameters. In this implementation, we construct an RB-SVM by using an RBF as the kernel function (Eq. 14).

Finally, note that although the SVM classifiers described above are binary classifiers, they are easily combined to handle the multiclass case. A simple, effective combination trains N one-versus-rest classifiers (say, "one" positive, "rest" negative) for the N-class case and takes the class for a test point to be that corresponding to the largest positive distance [26].

D_2. Discriminant Analysis (QDA and LDA): It provides an optimal classification rule (in the sense of minimising known errors) for discriminating of one population against another. Given a training sample consisting of m alternatives whose classification is *a priori* known, the objective of the method is to develop a set of discriminant functions maximizing the ratio of among-groups to within-groups variance. In our application, both linear (LDA) and quadratic discrimininant analysis (QDA) are applied. QDA can be seen as an extension of LDA allowing for curved (instead of linear) boundaries between populations. In the general case where the classification involves q groups, *q-1* linear/quadratic functions of the following forms are used:

$$Z_{kl} = a_{kl} + b_{kl_1} g_1 + b_{kl_2} g_2 + \ldots + b_{kl_n} g_n, \tag{16}$$

and

$$Z_{kl} = a_{kl} + \sum_{i=1}^{n} b_{kl_i} g_i + \sum_{i=1}^{n} \sum_{h=1}^{n} c_{kl_{ih}} g_i g_h \tag{17}$$

for the linear and the quadratic case, respectively. The calculation of the coefficients b_{kl_i}, $c_{kl_{ih}}$ and the constant term a_{kl} is achieved using the within-group covariance matrices for classes C_k and C_l. Classification is achieved using the score of an alternative on each discriminant function.

Results

For the evaluation of our methodology, EEGs from 25 subjects (13 epileptic and 12 normal) from the Neurology Department at University Hospital of Ioannina, Greece is used. More precisely, our dataset consists of 858 prototype transient events (274 ES, 254 EMG, 81 EOG, and 249 SAA) annotated by two experienced neurologists. The training set consists of 50% selected transient events from each category, and the remaining transient events are used for testing. The performance of our methodology is determined measuring the Sensitivity (*Se*), the Selectivity (*Sel*). Also we used the Accuracy (*Acc)* which is defined as:

$$Acc = \frac{\#of\ correctly\ classified\ events}{total\ \#\ of\ classified\ events}. \tag{18}$$

Table 1 displays the results obtained from the use of three classifiers (SVM, QDA, and LDA). All the tested

Table 1: Performance of the SVM, QDA and LDA classifier

	SVM		**QDA**		**LDA**	
	Se **(%)**	***Sel*** **(%)**	***Se*** **(%)**	***Sel*** **(%)**	***Se*** **(%)**	***Sel*** **(%)**
ES	85.42	88.81	67.15	82.88	63.5	86.14
EMG	85.77	80.71	89.06	83.21	85.83	80.15
EOG	58.82	72.29	82.5	80.49	82.93	87.18
SAA	94.74	88.89	79.03	70	85.48	69.28
Acc **(%)**	**84.83**		**78.55**		**78.32**	

classification schemes performed comparably but the use of SVM was found to be the most efficient. The transient event classification methodology was tested on our EEG dataset and demonstrated a *Se* and *Sel* of 85.42%, 88.81% for ESs, 85.77%, 80.71% for EMG, 58.82%, 72.29% for EOG, and 94.74%, 88.89% for SAA, respectively.

Discussion

In this study, we introduce an innovative method which detects transient events in EEG recordings and classifies those as ESs, EMG, EOG and SAA. Our method accomplishes signal preprocessing and transient event detection, clustering of transient events, feature extraction and classification of transient events by means of SVMs. The use of SVM is advantageous since it improves the efficiency of the methodology. This can be confirmed comparing with the QDA and LDA classifiers. SVM demonstrated a 6% increase in performance (the classification accuracy for the QDA and LDA was around 79%, Fig. 5).

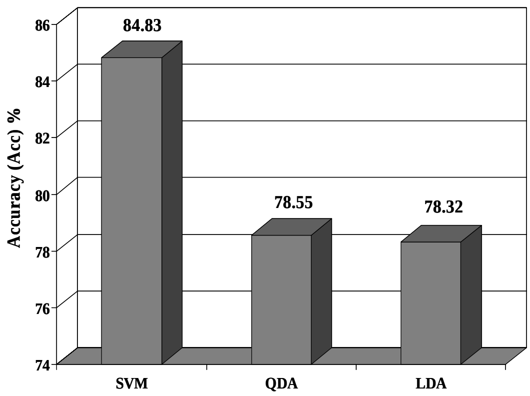

Figure 5: Comparison of classification accuracy for the SVM, QDA and LDA classifier

Looking at a more technical level inside the SVM, it should be noted that the selection of the kernel *K* is of major importance for the performance of the classifier. In our case an RBF kernel (Eq. 14) has been applied. Alternative approaches, such as linear (Eq. 12) or polynomial (Eq. 13) were not used due to the nature of our problem. The linear kernel cannot handle non linear separable problems; the polynomial kernel has more hyperparameters than the RBF kernel, fact that influences the complexity for model selection decision [26,27].

The literature presents various approaches which applied to EEG classification. The majority of the related works adressed only ES detection where the number of false positives due to spike-like artefacts and background activity is high. Therefore, comparison of our methodology with other detection techniques given in the literature is quite difficult. Furthermore, a direct comparison with other methods is difficult due to the data sources used, due to different recording types, displaying montages, channel numbers, degree of artefact and status of subject used.

It is well established that, apart from the ES detection on a single channel itself, other contextual information is also vital to the neurologists when classifying an event as epileptic or non-epileptic. This information is related to other channels ES activity which takes place at the same time. The proposed method does not take advantage of the spatial information but "inspects" each recording channel indivindually. Our future work will focus on the use of such information in making the final diagnosis.

Conclusions

An innovative approach for the automatic detection and classification of transient events in multichannel EEG recordings is presented. Our methodology does not eliminate the artefacts (EMG, EOG) and the background activity (SAA), but it detects and classifies it into the appropriate categories successfully using an SVM classifier. The classification of every transient event is a task that has never been accomplished before succesfully. The proposed methodology can be used as an assistant to the neurologists in making their decision during clinical practice. However, further testing and clinical evaluation is required.

References

[1] SUNDARAM M., SADLER R.M., YOUNG G.B. and PILLAY N. (1999): 'EEG in Epilepsy: Current Perspectives', *Can. J. Neurol. Sci.,* **26**, pp. 255-262

[2] GOTMAN J. and GLOOR P. (1976): 'Automatic recognition of inter-ictal epileptic activity in the human scalp EEG recordings', *Electroenceph. Clin. Neurophysiol.* , **41**, pp. 513-529

[3] TARASSENKO L., KHAN Y.U and HOLT M.R.G. (1998): 'Identification of inter-ictal spikes in the EEG using neural network analysis', *IEE Proc Sci. Meas. Technol.*, **145**, pp. 270-278

[4] WILSON S.B. and EMERSON R. (2002): 'Spike Detection: a review and comparison of algorithms', *Clin. Neurophysiol.*, **113**, pp 1873-1881

[5] PARK H.S., LEE Y.H., KIM N.G., LEE D.S. and KIM S.I. (1998): 'Detection of epileptiform activities in the EEG using neural network and expert system', *Medinfo* , **9**, pp. 1255–1259

[6] ACIR N., OZTURA I., KUNTALP M., BALKAN B. and GUZELIS C. (2005): 'Automatic detection of Epileptiform Events in EEG by a Three Stage Procedure Based on Artificial Neural Networks', *IEEE Trans. on Biomed. Eng.,* **52**, pp. 30-40

[7] LIU H.S., ZHANG T. and YANG F.S. (2002): 'A Multistage, Multimethod Approach for Automatic Detection and Classification of Epileptiform EEG', *IEEE Trans. on Biomed. Eng.,***49,** pp. 1557-1566

[8] GOTMAN J. (1982): 'Automatic recognition of epileptic seizures in the EEG', *Electroenceph. Clin. Neurophysiol.*, **54**, pp. 530–540

[9] KTONAS P., LUOH W., KEJARIWAL M. and SEWARD M. (1981): 'Computer-aided quantification of EEG spike and sharp characteristics', *Electroenceph. Clin. Neurophysiol.*, **51**, pp. 237-243

[10] KALAYCI T. and OZDAMAR O. (1995): 'Wavelet pre-processing for automated neural network detection of EEG spikes', *IEEE Eng. Med. Biol.*, **14**, pp. 160-166

[11] LOPES F.H. DA SILVA (1975): 'Detection of nonstationarities in EEGs using the autoregressive model-an application to EEGs of epileptics', in Dolce G. Kunkel H. (Ed), 'CEAN: computerized EEG analysis', (Stuttgart, Germany), pp. 180-199

[12] WEBBER W.R., LITT B., WILSON K. and LESSER R.P. (1994): 'Practical detection of epileptiform discharges (ED's) in the EEG using an artificial neural network: a comparison of raw and parameterized EEG data', *Electroenceph. Clin. Neurophysiol.* **91**, pp. 194-204

[13] JAMES C.J., JONES R.D., BONES P.J. and CARROLL G.J. (1999): 'Detection of epileptiform discharges in the EEG by a hybrid system compromising mimetic, self-organized artificial neural network, and fuzzy logic stages', *Clin. Neurophysiol.* **110**, pp. 2049-2063

[14] PANG C.C.C., UPTON A.R.M., SHINE G. and KAMATH M.V. (2003): 'A Comparison of Algorithms for Detection of Spikes in the Electroencephalogram', *IEEE Trans. on Biomed. Eng.,* **50**, pp. 521-525

[15] DAVEY B.L., FRIGHT W.R., CARROLL G.J. and JONES R.D. (1989): 'Expert system approach to detection of epileptiform activity in the EEG', *Med. Biol. Eng. Comput.*, **27**, pp. 365-370

[16] CASTELLARO C., FAVARO G., CASTELLARO A., CASAGRANDE A., CASTELLARO S., PUTHENPARAMPIL D.V. and SALIMBENI C.F. (2002): 'An artificial intelligence approach to classify and analyse EEG traces', *Neurophysiol. Clin.,* **32**, pp. 193-214

[17] DINGLE A.A, JONES R.D, CARROLL G.J. and FRIGHT W.R., (1993): 'A multistage system to detect epileptiform activity in the EEG', *IEEE Trans. Biomed. Eng.,* **40,** pp. 1260-1268

[18] TZALLAS A.T., KATSIS C.D., KARVELIS P.S., FOTIADIS D.I., KONITSIOTIS S. and GIANNOPOULOS S. (2004): 'Classification of Transient Events in EEG Recordings', Proc. of the IEE Medical Signal and Information Processing Conference (MEDSIP), Sliema, Malta, 2004, p. 14-20

[19] EXARCHOS T.P., TZALLAS A.T., FOTIADIS D.I., KONITSIOTIS S. and GIANNOPOULOS S. (2005): 'A Data Miming Based Approach for the EEG Transient Event Detection and Classification', Proc. of the 18th IEEE International Symposium on Computer-Based Medical Systems (CBMS), Dublin, Ireland, 2005, p. 14-20

[20] WAHLBERG P. and LANTZ G. (2000): 'Methods for robust clustering of epileptic EEG spikes', *IEEE Trans. Biomed. Eng.,* **47**, pp. 857-868

[21] KOTHARI R. and PITTS D. (1999): 'On finding the number of clusters', *Patt. Rec. Letters*, **20**, pp. 405-416

[22] CHRISTIANINI N. and SHAWE-TAYLOR J. (2000): 'An Introduction to Support Vector Machines', (Cambridge University Press)

[23] BURGES C.J.C. (1998): 'A tutorial on Support Vector Machines for Pattern Recognition', in Data Mining and Knowledge Discovery, Fayyad U. (Ed), (Kluwer Academic Publishers, Boston), pp. 121-167

[24] FISHER R.A. (1936): 'The use of multiple measurements in taxonomic problems', *Annals of Eugenics,* **7**, pp. 179-188

[25] ALTMAN E.I., AVERY R., EISENBEIS R. and STINKEY J. (1981): 'Application of Classification Techniques in Business, Banking and Finance', *Contemp. Stud. Econ. Financ. Anal.*, **3**

[26] CORTES C. and VAPNIK V. (1995): 'Support-vector network', *Mach. Learn.,* **32**, pp. 273-297

[27] JOACHIMS T. (1998): 'Making large-scale SVM learning practical', in Scholkopf B., Burges C.J.C, Smola A.J (Ed), 'Advances in kernel methods – support vector learning', (MA: MIT Press, Cambridge)

ISSN: 1727-1983 © 2005 IFMBE