International Journal of Medical, Medicine and Health Sciences
ISSN: 2517-9969
Vol:1, No:2, 2007

Classification Control for Discrimination between Interictal Epileptic and Non – Epileptic Pathological EEG Events

Sozon H. Papavlasopoulos, Marios S. Poulos, George D. Bokos, and Angelos M. Evangelou

***Abstract*—**In this study, the problem of discriminating between interictal epileptic and non- epileptic pathological EEG cases, which present episodic loss of consciousness, investigated. We verify the accuracy of the feature extraction method of autocross-correlated coefficients which extracted and studied in previous study. For this purpose we used in one hand a suitable constructed artificial supervised LVQ1 neural network and in other a cross-correlation technique. To enforce the above verification we used a statistical procedure which based on a chi- square control. The classification and the statistical results showed that the proposed feature extraction is a significant accurate method for diagnostic discrimination cases between interictal and non-interictal EEG events and specifically the classification procedure showed that the LVQ neural method is superior than the cross-correlation one.

***Keywords*—**Cross-Correlation Methods, Diagnostic Test, Interictal Epileptic, LVQ1 neural network, Auto-Cross-Correlation Methods, chi-square test.

I. Introduction

It is known that determining whether a person with "seizures", "spells" or other episodic unusual behaviour, actually has epilepsy presents difficulties. For example episodic loss of consciousness need not signal epilepsy but could result from loss of blood supply to the brain from diseases of the blood vessels or the heart itself. Periodic low blood sugar and certain types of migraine headache may also lead to loss of consciousness [1]. Therefore, Non-Epileptic Events (NEEs) may be due to different organic or non-organic disorders. The diagnosis of Non-Epileptic Attack Disorder (NEAD) involves both exclusion of organic causes of NEEs and elucidation of positive phenomena of this entity [2]. The distinct entity of NEAD does not allude to any specific psychological mechanism and this term includes a variety of synonyms like Pseudo Epileptic Seizure (PES), psychogenic seizure, pseudo seizure, hysterical seizure, hystero-epilepsy and functional seizure. The subject has recently attained renewed interest as intensive monitoring has diagnosed many cases of refractory seizures (20% or more) as non-epileptic seizures [3]. In the case of Epileptic events, the condition where the brain itself is the cause of periodic spells, the classic diagnostic approach has always been to perform an EEG and search for epileptiform "spikes" or "spike and waves" which may signify epilepsy [4, 5]. Electroencephalography remains a major complex technique in differentiating epilepsy and non-epileptic attacks like NEAD, syncope, narcolepsy, cataplexy, sleep disorders, etc. Proper clinical history and observation of an attack may not be sufficient for diagnosis and, therefore, ictal and postictal EEG, 24 hours ambulatory EEG and video EEG can be of immense help for the purpose. Long term monitoring (LTM) for epilepsy is the technological advancement to improve the yield of EEG data in differentiating Epileptic Seizure (ES) from Non- Epileptic Seizure (NES). LTM includes radio telemetry, cable telemetry and cassette recorders [6]. Suggestion and induction techniques along with simultaneous continuous video-EEG monitoring have been used to differentiate between EE and NEE. These include iv saline infusion, alcohol patch technique and hypnosis and NEEs could be induced in 77-82% cases [7, 8].

Ideally, an EEG is performed during an actual clinical or "ictal" event during which time runs of epileptiform discharges would be expected. However, ictal events may be few and far between. In practice most epileptics demonstrate epileptiform activity even in-between seizures (interictally). The human eye is the "gold standard" for recognizing epileptiform activity and to distinguish it from artifactual signals and from EEG activity that may mimic epileptiform activity but is benign ("normal variants"). However the unaided human eye cannot efficiently distinguish the specific details of interictal epileptic activity that are valuable regarding a final epileptic diagnosis [4, 5].

Our study [9], a diagnostic testing method used to discriminate between interictal epileptic EEG and non-epileptic pathological EEG events, this method based purely on signal processing and describes an algorithm which is

Manuscript received February 4, 2006.

S. Papavlasopoulos, post-graduate, Department of Archives and Library Science, Ionian University. Palaia Anaktora PO Box 96 49100 Corfu Greece (e-mail: sozon@ionio.gr).

M. Poulos, Phd, Adjunct Teacher Department of Archives and Library Science, Ionian University. Palaia Anaktora PO Box 96 49100 Corfu Greece (e-mail: mpoulos@ionio.gr http://www.ionio.gr/~mpoulos).

G. Bokos, Full Professor, Department of Archives and Library Science, Ionian University. Palaia Anaktora PO Box 96 49100 Corfu Greece (e-mail: gbokos@ionio.gr http://www.ionio.gr/~gbokos).

A. Evangelou, Department of Exp. Physiology, School of Medicine, University of Ioannina. University Campus of Ioannina 45110 Greece (e-mail: evagel@uoi.gr).

based on the estimation of a number of auto-correlated coefficients extracted from an interictal epileptic EEG segment. In particular, these coefficients are correlated with the coefficients of EEG segments of epileptic and non-epileptic cases. Finally, the auto-correlation coefficients are extracted in a particular spectrum in contrast to the traditional methods where the final diagnosis of epilepsy depends on searching for epileptiform "spikes" or "spike and waves" [4, 5]. In this way the autocorrelation coefficients of a specific interictal epileptic EEG segment may be used as a pattern recognition tool for epileptic diagnosis.

The statistical results of this method corroborate the previous research [11, 12], in that it is possible to correlate alpha, beta and gamma activities with epileptic activity. This conclusion is justified because in the experimental part the selected spectrum of each EEG segment that participated in the proposed method contained dominant alpha activity and fewer beta and gamma activities, which, however, influenced significantly the results, using cross-correlation technique.

Furthermore, the same problem was investigated with the same feature data using an artificial LVQ1 neural network [10]. This experiment took place in order to be corroborated the results of our previous work [9]. The results of this experiment [10] showed that the LVQ1 neural network classified better in the frequency 8-40 Hz than the latest study [9], which yielded best results in the area between 5 - 70 Hz. Specifically, the two categorization methods didn't yield equivalent results in the same frequencies that experimentally referred.

The aim of the present study is to statistically compare the above results of the above methods in order to extract accurate statistically conclusion for the features of the proposed feature extraction mechanism. In order to evaluate the statistical significance of the classification scores obtained in the experimental part, the chi-square test was applied to the results. The two feature extraction methods presented are also compared in terms of their Cramer coefficient of mean square contingency, φ. Results are seen to be statistically significant at the a = 99.5% level of significance.

So in the present work we compare statistically all the results of the both methods in order to extract the following two ascertainments:

1. The hypothesis that the dominant alpha activity and fewer beta and gamma activities carried interictal epileptic features is fundamental statistically.
2. The statistically improvement of the validation of the feature extraction method [13].

For the implementation of this we used statistically tools in a similar difficult problem [13].

In more details, we used a good fitted chi square model in order to compare the both results regarding a null Hypothesis, which indicates appropriately the classification problem. Furthermore, we made more robust this evaluation using the Cramer coefficient [13] of mean square contingency.

II. The Feature Extraction Method

A. The Basis of the Algorithm

This study is based on the *hypothesis that the shape of a segment of an EEG signal may be described by the degree of asymmetry around a characteristic point* [9]. The degree of asymmetry of a segment is obtained via the Pearson criterion [13] and is described by the following equation:

$$S = \frac{\bar{X} - M_o}{s} \quad (1)$$

where:

- S the degree of asymmetry,
- $\bar{X}$ the mean value of the signal segment,
- M_o the value of the characteristic signal (data) point,
- s the standard deviation of a signal segment.

The degree of asymmetry may be characterized as an appropriately fit index because it includes all the necessary characteristics of the EEG for our purpose. In other words, using the highest peak of the spectral density of an EEG as a symmetric axis, the extracted index in the interictal epileptic case features the positions of the waves and spikes. Moreover, it also carries the general feature of the distribution across the spectrum. This consideration may be characterized as innovative because it is possible to detect characteristic differences between pathological cases that yield similar EEG recordings such as those referred to in the introduction. Furthermore, in the present study we chose to extend our research to include gamma activity, 5-70 Hz (fig. 1).

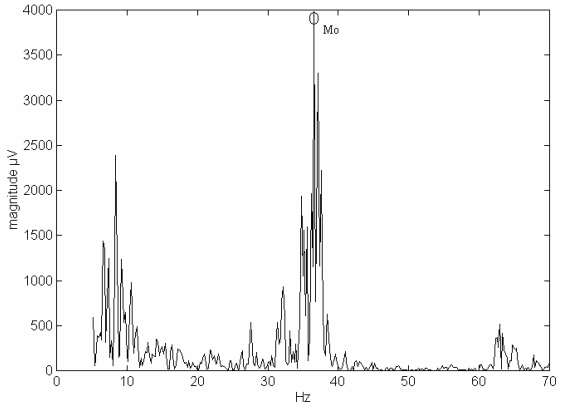

Fig. 1 Spectral analysis of an EEG signal. Only the sub-band (5-70 Hz) is shown

B. The Autocorrelation Coefficients

In this study we considered that the original EEG signal *x(n)* was segmented into k sequential overlap segments. Thus, we created *sequence* (w_k),

where:

$$\{w_k\} = \{x_{1+hk}, x_{1+hk+1,\ldots,} x_{1+2hk+f}\} \quad (2)$$

And $k = 0,1,2,...,\frac{n}{2h}$ *with f* and *h* the constants described in the experimental part.

In our case we considered that each EEG segment $x(w_k)$ with a length of *N* was partitioned into k non-overlapping sequences with a length of L so that *kL=N*. The *k* non-overlapping sequences can be expressed as:

$$x_m(w_k) = x(w_k + mL), w_k = 0,1,...,L-1 \text{ and}$$
$$m = 0,1,...,k-1.$$

Thereinafter the Power Spectral Density $\hat{P}_B$ of each EEG overlap segment $x_m(w_k)$ was computed using Bartlett's periodogram method [14] as follows:

$$\hat{P}_B(e^{j\omega}) = \frac{1}{N}\sum_{m=0}^{k-1}\left|\sum_{w=0}^{L-1} x(w_k + mL)e^{j\omega}\right|^2 \quad (3)$$

It is considered that the Bartlett estimate $\hat{P}_B(e^{j\omega})$ is an asymptotically unbiased and consistent estimate of the power spectrum $\hat{P}_x(e^{j\omega})$.

Furthermore, we considered that a sequence of frequencies {f} = {f1, f2, f3,…, fn} is the same length as set $\hat{P}_B$, the values of f_1 and f_n being determined in the experimental part. Then, if fg is the element of {f} sequence which corresponds to the P_g element of the {*P*} sequence,

$$\text{Where: } P_g = \max(\hat{P}_B(e^{i\omega})) \text{ and } 1 \le g \le n \quad (4)$$

Then equation (1), taking into account equations (3, 4), is modified as follows:

$$S_k = \frac{\bar{\hat{f}} - f_g}{\sqrt{\frac{\sum\left|\hat{P}_B(e^{i\omega})\right|^2 - \frac{\left|\sum \hat{P}_B(e^{i\omega})\right|^2}{N}}{N-1}}} \quad (5)$$

where $\hat{P}_B$ is given by equation 3.

Thereinafter, we considered set {D} of sequences which consist of the following:

$$\{D\} = \{\hat{D}_1, \hat{D}_2, ..., \hat{D}_{k-1}\}, \text{ where:}$$

$$\hat{D}_1 = \{S_1, S_2\}, \hat{D}_2 = \{S_1, S_2, S_3\}, \hat{D}_3 = \{S_1, S_2, S_3, S_4\}, ..., \hat{D}_{k1} = \{S_1, S_2, S_3, ..., S_k\}.$$

Then the autocorrelation coefficients of the proposed method were computed as follows:

$$\hat{C} = [C_1, C_2, C_3, ...C_{K-1}] \quad (6)$$

where:

$$C_1 = \sqrt{\frac{\sum\left|\hat{D}_1\right|^2 - \frac{\left|\sum \hat{D}_1\right|^2}{N}}{N-1}}$$

In conclusion, these extracted autocorrelation coefficients may be characterized as a mapping of the variation of spectral density of an EEG as can be seen in figure 2 below.

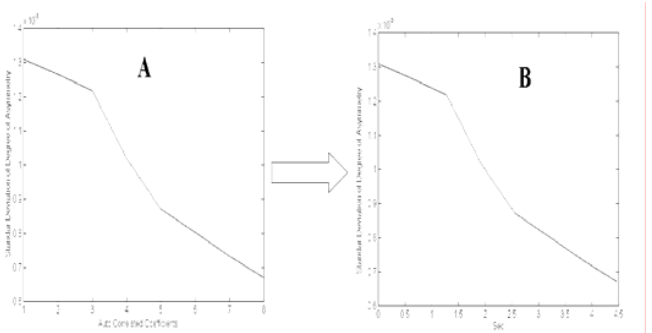

Fig. 2 In figure ***A*** an example of the variation of the autocorrelation coefficients is presented. In figure ***B***, figure ***A*** is transformed in order to present the variation of the standard deviation in relation to time

III. Classification Procedure via Cross-Correlation Method

In this stage the extracted set of auto-correlation coefficients $\hat{C}_x$ of an interictal epileptic EEG case were submitted to the cross-correlation [15] along with another set $\hat{C}_y$ (interictal epileptic or non-epileptic case) as described below:

$$r = \frac{\sum_{i=1}^{k-1}(\hat{C}x_i - \bar{\hat{C}}x_i)(\hat{C}y_i - \bar{\hat{C}}y_i)}{\sqrt{\sum_{i=1}^{k-1}(\hat{C}x_i - \bar{\hat{C}}x_i)^2 \sum_{i=1}^{k-1}(\hat{C}y_i - \bar{\hat{C}}y_i)^2}} \quad (7)$$

The extracted cross-correlation coefficient is a number between -1 and 1, which measures the degree to which two variable sets are linearly related. In our study we considered that the auto-correlated unknown EEG set has a perfect positive linear relationship with the auto-correlated interictal epileptic set $\hat{C}_x$ when the cross-correlation coefficient is approximately 1.

IV. Classification Procedure via LVQ Neural Network

The extracted set of auto-correlation coefficients used as feature vectors for classification. These vectors (codebook) are fed into an LVQ1 classifier, [16], first for training and then for the actual classification of unknown input vectors. During the training process, the codebook vectors directed towards the data vectors of the same class and distanced from those codebook vectors of a different class. The adaptation of the weights of the neurons carried out iteratively, based on the Euclidean distance measure. Specifically, the architecture of the LVQ1 network used to classify the epileptic or non-

epileptic feature vectors (p = 8). Input vectors of dimensionality 8 x 1 are weighted and fed to the first layer of neurons, known as the competitive layer. These neurons compete for inputs in a "greedy" way; hence the layer name. Four (4) such neurons form the competitive layer in our case. The output of the competitive layer, which is in fact a grouping of the inputs into subclasses, fed to the second linear layer, which groups subclasses into target classes. The weights connecting the two layers take on binary values of zero or one, indicating mere class membership and not actual weighting.

V. Experimental Part

In this study two (2) data types were recorded. On one hand 42 interictal epileptic EEGs from diagnosed epileptic individuals were recorded and on the other hand 44 EEGs from diagnosed pathological cases, who had presented loss of consciousness, were also recorded. It must be noted that for all the EEGs of both data types it was impossible to diagnose with the eye or with known computer methods based on detection using "spikes" or "spike and waves". That is because in the interictal epileptic cases we selected original EEG segments which were devoid of characteristic epilepticform spikes, (see figure 3). Furthermore, the epileptic and non-epileptic EEG segments belonged to different adult individuals.

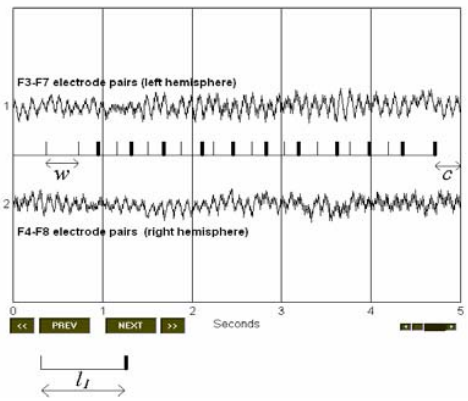

Fig. 3 An example of the overlapping segmentation of two EEGs of 5 sec duration each, where $\mathbf{l_1}$ is the selected length of each EEG segment, **w** is the overlap window and **c** is the non-processed EEG segment. In case 1 there is an EEG segment which sources from F3-F4 electrode pairs(left hemisphere) while in case 2 there is an EEG segment which sources from F4-F8 electrode pairs (right hemisphere). In both cases the gamma activity is evident

All recordings were taken using a digital electroencephalograph with RHY-100 Stellate software. Subjects were at rest, with closed eyes. Voltage difference (in μVolts) was recorded between leads O2 and CZ. The selection of these leads is justified because it is known that from these regions of the scalp can be extracted those faster activities such as alpha, beta and gamma. All EEG recordings lasted for twenty (20) continuous seconds (the duration having been selected after experimentation), thus producing a 4000 samples long record each at a 200 Hz sampling rate. Further processing was carried out off-line, in Matlab 5.2, on a Pentium PC. Furthermore, for the extraction of characteristic interictal epileptic set $\hat{C}_x$, an interictal epilepticform EEG was recorded lasting three (3) continuous minutes, thus producing a 36000 samples long record.

These data were submitted into two classification methods, the cross- correlation method and the LVQ neural method. The scores of the classification procedures of these methods are presented below:

VI. The Feature Extraction

All (42+44=86) EEGs recorded were submitted to the auto-correlation procedure as described in section II.B. We ascertained that the best results were extracted using n=2000, F=1000, h=128 and k=8 in equation 2. Furthermore, after experimentation we adopted those frequencies between 5 and 70 HZ as the most suitable spectrum for each EEG segment. In total 86 auto-correlation vectors were yielded, each of size (1x8). In table II an example of two characteristic vectors of the extracted auto-correlation coefficients can be seen. After that, the complete database of EEGs (epileptic and non epileptic) was submitted to the auto-correlation procedure and thereafter each of the extracted auto-correlated sets was cross-correlated with the interictal epileptic set $\hat{C}_x$ (section III) in order to ascertain their degree of linear relationship. For the extraction of the interictal epileptic set the determination of the parameters of equation (8) took place as follows:

$$w=200,\ l=36000,\ l_1=4000,\ g=180,$$
$$\text{thus } c=l-gw=36000-(180*200)=0$$

TABLE I
AN EXAMPLE OF TWO AUTO-CORRELATION COEFFICIENT SETS

Auto-Correlation Coefficients *(C)*	
Interictal Epileptic EEG	Non- Epileptic EEG
0.0280	0.0059
0.0238	0.0021
0.0192	0.0025
0.0169	0.0023
0.0146	0.0024
0.0117	0.0021
0.0105	0.0016
0.0091	0.0017

Hence, for this extraction 180 auto -correlated sets of coefficients were produced from the original epilepticform EEG and the appropriate set $\hat{C}_x$ was selected according to section III.

VII. The Results of the Cross-Correlated Method

A. The Selection of the characteristic interictal epileptic set $\hat{C}_x$

The selection is based on the claim that a particular characteristic segment of an epileptic EEG may carry specific epileptiform features [17]. For the determination of this segment we used a method that is based on the overlapping segmentation of the original EEG. The algorithm of this method is described as follows:

1. An original EEG segment *x(n)* of length *l* was segmented into *g* overlap segments of lengths l_1 with overlap window *w*. It must be noted that the determination of these lengths *(l, l_1, w)* was based on previous studies [12] and this was corroborated in the experimental part of the present study. The correlation of these lengths is determined as follows (see figure 3):

$$gw + c = l, \quad \text{where:} \quad 0 \le c \le l_1 \tag{8}$$

It should be noted that the values of parameters *g, w* and l_1 are determined in the experimental part and depend on EEG recording conditions such as the sampling rate, the duration of recording and the adapted filters. For better comprehension in our experiment we adopted as the most suitable values for the above parameters: *w=200, l=36000, l_1=4000, g=180, thus c=l-gw=36000-180*200=0,* these values are also mentioned in the experimental part.

2. Thereinafter the selected EEG segments were submitted to the autocorrelation procedure as described in section II.B.

3. Then the extracted sets of auto-correlated coefficients $\{\hat{C}_1, \hat{C}_2, \hat{C}_3, ...\hat{C}_g\}$ were cross-correlated with the first set $\hat{C}_1$ according to section III. Hence, a set of $\hat{r}_g$ cross-correlated coefficients was produced,

where: $\hat{r}_{1g} = \left[r_{11}, r_{12}, ..., r_{1g} \right]$

4. The value $\overline{\hat{r}}_{1g}$ was calculated.

5. Finally, from the above set $\hat{r}_{1g}$, the coefficient that was nearest in value to $\overline{\hat{r}}_{1g}$ was selected as the ideal cross-correlated r_x coefficient. This meant that selected set $\hat{C}_x$, when cross-correlated with the other sets, yielded a new set $\hat{r}_{xg}$ of cross correlated coefficients with the best linear relationship,

where: $\hat{r}_{xg} = \left[r_{x1}, r_{x2}, ..., r_{xg} \right]$

For this reason estimated set $\hat{C}_x$ may be characterized as ideal for our purpose.

B. The extraction of the cross-correlation coefficients

In table II the results of the cross-correlation coefficients, which were extracted according to the processing procedure described in section VII.A, are presented.

In more details, this table shows that the cross-correlated coefficients in the interictal epileptic EEG case range between 0.80 and 0.99 while in the non-epileptic case they range between 0.05 and 0.90. The first general conclusion to be drawn is that the interictal epileptic sets present a better correlation with the interictal epileptic set $\hat{C}_1$ than the non epileptic sets because their values are nearer the unit.

TABLE II
THE EXTRACTED CROSS CORRELATION COEFFICIENTS r

Cross-Correlation Coefficients (*r*)							
Interictal Epileptic EEG				Non -Epileptic EEG			
1	0.93	24	0.99	1	0.67	24	0.78
2	0.95	25	0.91	2	0.84	25	0.72
3	0.98	26	0.86	3	0.08	26	0.68
4	0.96	27	0.93	4	0.67	27	0.27
5	0.93	28	0.90	5	0.80	28	-0.14
6	0.89	29	0.93	6	0.34	29	0.68
7	0.90	30	0.94	7	0.87	30	-0.78
8	0.99	31	0.95	8	0.23	31	0.89
9	0.82	32	0.96	9	0.47	32	0.91
10	0.99	33	0.91	10	0.13	33	-0.09
11	0.96	34	0.85	11	0.71	34	0.85
12	0.98	35	0.94	12	0.13	35	0.15
13	0.89	36	0.95	13	0.71	36	0.37
14	0.86	37	0.93	14	-0.05	37	0.11
15	0.80	38	0.97	15	-0.33	38	0.72
16	0.83	39	0.92	16	0.24	39	0.78
17	0.98	40	0.81	17	0.08	40	-0.05
18	0.93	41	0.94	18	0.90	41	0.88
19	0.92	42	0.88	19	0.63	42	0.56
20	0.95	43		20	-0.56	43	0.87
21	0.94	44		21	0.59	44	0.24
22	0.91	45		22	0.67	45	
23	0.90	46		23	-0.21	46	

C. Tests of (least-squares) Correlation Coefficients

In this stage we controlled the value of the r coefficient, which considered as significant for the classification decision. In such tests, *r* is the sample-derived estimate of ρ. Then we considered that the null hypothesis is: $H_0 : \rho_0 = 0$. Therefore, the sampling distribution of *r* for a population that has zero correlation $(\rho = 0)$ has a mean value of $\mu = 0$ and. Hence, a t-statistic can be calculated as: $\sigma = \sqrt{\frac{(1-r^2)}{k-2}}$

$$t = \frac{r-\mu}{\sigma} = \frac{r}{\sqrt{\frac{(1-r2)}{k-2}}} = \frac{r\sqrt{k-2}}{\sqrt{1-r^2}} \quad (9)$$

The next step was to determine the appropriate value of the r coefficient in order to characterize it as a significant linear relationship between the correlated sets in our experiment. Thus, having $k=8$, and the degree of freedom $v=k-2=6$ we chose $a=0.01$ and thus found critical $t_{\alpha/2}=3.707$. Then the significant value of r was calculated as follows:

$$t_{\alpha/2} = \frac{r\sqrt{k-2}}{\sqrt{1-r^2}} \Rightarrow 3.707^2 = \frac{6r^2}{1-r^2} \Rightarrow r = \pm 0.83$$

In conclusion, in our case, coefficient r may be characterized as significant when the null hypothesis is rejected ($1 \le |r| \le 0.83$). Taking this into account, table II is modified to table III.

TABLE III
THE CLASSIFICATION SCORE VIA CROSS-CORRELATION COEFFICIENTS

Cross-Correlation Coefficients (r)		
CATEGORIES	TRUE	FALSE
INTERICTAL-EPILEPTIC	38/42=90%	4/42=10%
NON-EPILEPTIC	36/44=82%	8/44=18%

VIII. The Results of the LVQ1 (Method)

In this procedure, we trained LVQ1 neural networks for each case. In order to differentiate class A (epileptic) from class B (non-epileptic), twenty (22) feature vectors (11 for each class) are used for training the classifiers and the other forty-four (44) are used for testing. The LVQ1 neural network, which was used in the aforementioned training procedure, is described in sections II and IV and was trained for a total of 500 cycles (epochs) with a learning rate in the order of 0.7 (fig 4). The results of this classification are presented in table IV.

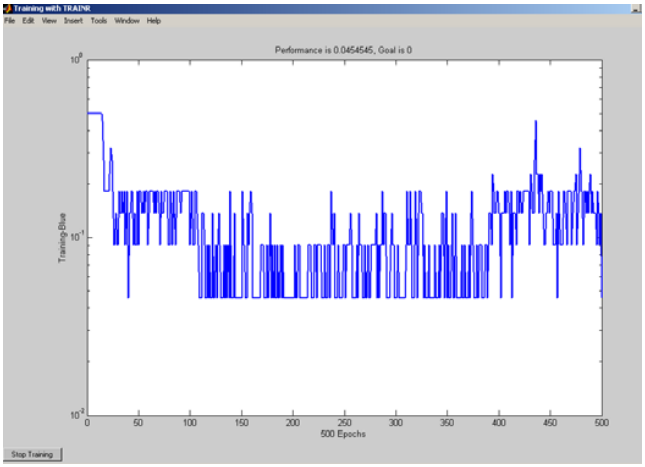

Fig. 4 Error plot while training an LVQ1 network using EEG feature vectors

TABLE IV
THE CLASSIFICATION SCORE VIA LVQ NEURAL NETWORK

Cross-Correlation Coefficients (r)		
CATEGORIES	TRUE	FALSE
INTERICTAL-EPILEPTIC	36/42=86%	6/42=14%
NON-EPILEPTIC	39/44=89%	5/44=11%

IX. The Statistical Evaluation of These Results

The statistical procedure of both classification methods is based on the statistical x square control of the results. Thus, for implementation of this we used the following statistical controls.

The results of a classification experiment can be put into a two-ways contingency table, [3,4]. A two-ways contingency table is structured on the basis of two criteria, along its two dimensions. Here we use "subject belongs to class i" (epileptic) as the first criterion (vertical dimension) and "subject is classified into class j"(non-epileptic) as the second criterion (horizontal dimension). An ideal classification method should produce a diagonal matrix of classification scores ("subject belongs to class i"(epileptic) and "subject is classified into class i"(epileptic)), corresponding to full dependency between the two above criteria, while practical methods would tend to this behavior. Evaluation of the statistical significance of the classification results is thus transformed into a hypothesis testing problem: The null hypothesis of independence of the two criteria is tested against the alternative hypothesis of dependence. The test statistic used for this purpose is the χ^2. Statistically significant classification results correspond to rejection of the null hypothesis at a satisfactory level of significance.

Let the contingency matrix S be of dimensions (r x c), meaning r rows and c columns, and let the (i, j)-th entry of S, $\{S(i, j) = f_{ij}; i = 1, \ldots, r; j = 1, 2, \ldots, c\}$ denote *observed* frequency of occurrence of the event (i, j) ("subject belongs to class i" and "is classified into class j") and $\{e_{ij}; i = 1, \ldots, r; j =$

1, 2, …, c} denote *expected* frequency of occurrence of the event (i, j). Then the test statistic is given by

$$\chi^2 = \sum_{i=1}^{r}\sum_{j=1}^{c} \frac{(f_{ij}-e_{ij})^2}{e_{ij}} \quad (10)$$

which asymptotically follows the χ^2 distribution with (r-1)(c-1) degrees of freedom. When unknown, expected frequencies can be estimated from S using

$$e_{ij} = \frac{R_i C_j}{N} \quad (11)$$

where N is the total number of events in S, R_i is the sum across the i-th row of S and C_j is the sum across the j-th column of S.

The degree of dependence between the two criteria can also be measured by the Cramer coefficient [13] of mean square contingency,

$$\phi = \sqrt{\frac{\chi^2}{N\min(r-1, c-1)}} \quad (12)$$

Coefficient φ takes on values between 0 (independence) and 1 (full dependence). Two classification methods can in fact be compared in terms of their Cramer coefficient, as to the statistical significance of their results. Note that for 2x2 contingency tables, (12) becomes

$$\varphi = \frac{\chi}{\sqrt{N}} \quad (13)$$

In the test of case, the results form 2 x 2 contingency table. Expected frequencies accompany observed frequencies in the cells of Tables III, IV. As it can be seen in Table V, the χ^2 values of the test statistic, as computed from the results in Table VI are [45,11], respectively, for the Cross-Correlated feature vectors and [47,63] for the Lvq1 feature vectors. As an example, for the correct positive classification cell (1,1) of Table V (Cross-Correlation features), χ^2 test value is computed as

$$x^2 = \frac{(38-22,47)^2}{22,47} + \frac{(4-19,53)^2}{19,53} + \frac{(8-23,53)^2}{23,53} + \frac{(36-20,47)^2}{20,47} = 45,11.$$

From the tables of the χ^2 distribution with one (1) degree of freedom and at the 99.5% level of significance, we obtain the critical value 7.879, which is lower than all test statistic values. The null hypothesis of independence is therefore rejected for the case and for both types of feature vectors. Furthermore, φ coefficient takes on value [0,26] for the experiment based on Cross-Correlation feature vectors and [0.28] for the Lvq1 feature vectors.

TABLE V
TEST CASE, SUBJECT EPILEPTIC VERSUS GROUP NON-EPILEPTIC CLASSIFICATION SCORES BASED ON LVQ1 AND CROSS-CORRELATION FEATURE VECTORS

	Lvq1		***Cross-cor feature vectors***		
Classified as: ⇒ Belongs to class: ⇓			epil	Non-epil	***Total***
epil	$\frac{36}{42}$(85,7%) *[20.02]*	$\frac{6}{42}$(24%) *[21,98]*	$\frac{38}{42}$(90,4%) *[22,47]*	$\frac{4}{42}$ (9, 6) *[19,53]*	42
Non-epil	$\frac{5}{44}$(11,4) % *[20,98]*	$\frac{39}{44}$(88,6%) *[23,02]*	$\frac{8}{44}$(18,2%) *[23.53]*	$\frac{36}{44}$(81,8% *[20,47]*	44
Total	41	45	46	40	86

TABLE VI
CHI-SQUARE TEST EVALUATION OF THE RESULTS IN TABLE IV
Critical χ^2 values from χ^2 tables (a = 0.995 level of significance) in brackets, along with the Cramer coefficient φ.

Features ⇒ Test cases: ⇓	**Lvq1**		**Cross-Correlated feature vectors**	
	χ^2 statistic		**χ^2 statistic**	**Probability**
	[$\chi^2_{(1, 0.995)}$]		**[$\chi^2_{(1, 0.995)}$]**	
	φ value		**φ value**	
Epil-Non-epil	47,63 [7.879] 0.28	P<0.001	45,11 [7.879] 0.26	P<0.001

X. Conclusions

In this study we investigated the reliability of diagnostic method the recognition of interictal epileptic and non-epileptic pathological EEG cases using auto-cross-correlation methods. For implementation of this we tested the vectors of the feature extraction via two known classification procedures. First the classic cross-correlation control against an artificial LVQ1 neural network. For more reliability of our experiment we used the aforementioned statistical techniques. In more details, from the tables of the χ^2 distribution with one (1) of freedom and at the 99.5% level of significance, we obtain the critical value 7.879, which is lower than all for both types of feature vectors (47.63, 45.11). The statistical results of the

proposed method corroborate the latest research [11] in that it is possible to correlate alpha, beta and gamma activities with epileptic activity. This conclusion is justified because in the experimental part the selected spectrum of each EEG segment that participated in the proposed method, contained dominant alpha activity and fewer beta and gamma activities, which, however, influenced significantly the results.

Furthermore, the experimental results of the spectral analysis show that the algorithm that is described in equation (5) corroborates the hypothesis that the shape of a segment of an EEG signal may be described by the degree of asymmetry around a characteristic point [9].

In conclusion, the classification and the statistical results showed that the proposed feature extraction is a significant accurate method for diagnostic discrimination cases between interictal and non-interictal EEG events and specifically the classification procedure showed that the LVQ neural method is superior than the cross-correlation one.

REFERENCES

[1] M.C. Brown, B.E. Levin, E. Ramsay, D.A. Katz, M.S. Duchowny, "Characteristics of patients with non-epileptic seizures," *J Epilepsy*, 1991, 4(5) pp. 225-229.

[2] H. Meierkord, R. Will, D.R. Fish, S.D. Shorvon, "The clinical features and prognosis of pseudoseizures diagnosed using video-EEG telemetry," *Neurology*, 1991, 41(10) pp. 1643-1646.

[3] V. Ramani, "Intensive monitoring of psychogenic seizures, aggression and dyscontrol syndromes," *Adv. Neurol,* 1986, 46(2), pp. 103-127.

[4] F.A. Gibbs, E.L. Gibbs, W.G. Lennox, "Electroencephalographic classification of epileptic patient and control subjects," *Arch Neural Psychiatric*, 1943, 50(2), pp. 111-128.

[5] V. Ramani, S. Whalen, R. Loewenson, "Ictal characteristics of pseudo-seizures," *Arch. Neurol.*, 1985, 42(9), pp. 1183-1187.

[6] C.D. Binnie, *Long-term monitoring, Comprehensive Epileptology*. New York: Raven Press, 1991, pp.88-110.

[7] J.J. Barry, O. Atzman, M.J. Morrell. "Discriminating between epileptic and non-epileptic events: the utility of hypnotic seizure induction," *Epilepsia* 2000, 41(1), pp. 81-84.

[8] T.S. Walczak, D.T. Williams, W. Berten, "Utility and reliability of placebo infusion in the evaluation of patients with seizures," *Neurology* 1994, 44(3), pp. 394-399.

[9] M. Poulos, F. Geogiacodis, V. Chrissicopoulos, A. Evangelou, "Diagnostic Test for the Discrimination between Interictal Epileptic and Non-Epileptic Pathological EEG Events using Auto-Cross-Correlation Methods," *American Journal of Electroneurodiagnostic Technology*, Dec 2003, v. 43, pp. 228-264.

[10] S. Papavlasopoulos, M. Poulos, A. Evangelou, "Feature Extraction from Interictal Epileptic and Non- Epileptic Pathological EEG Events for diagnostic Purposes using LVQ1 Neural Network," *Proceedings of seventh International Conference on Mathematics Methods in Scattering Theory and Biomedical Technology*, BIOTECH'7, 2005, Nimfaio, Greece.

[11] A. Medvedev, J.O. Willoughby, "Can hypersychronisation of fast (gamma) activity lead to generalized epilepticform discharges?" *Proceedings, Epilepsy Society of Australia*, 1999, 41.

[12] J.S. Barlow, "Methods of analysis of nonstationary EEGs with emphasis on segmentation techniques: a comparative review," *Clin. Neurophysiology*, 1985, 2(5) pp. 267 - 304.

[13] J.H. Zar, *Biostatistical Analysis*, New Jersey: Prentice-Hall, 1999, pp.72-73.

[14] S. Haukin, *Adaptive Filter Theory*, New Jersey: Prentice Hall, 1996 pp. 136-138.

[15] N. Morrison, F. Donald, *Multivariate Statistical Methods*, New York: McGraw-Hill Book Company, 1976, pp.128-130

[16] J. A. Kangas, T. Kohonen, J. T. Laakson, "Variants of Self-Organizing Maps," *IEEE Trans. Neural Networks*, 1990, 1:1, pp. 93-99.

[17] F. Mormann, K. Lehnertz, R.G. Andrzejak, C.E. Elger, "Characterizing preictal states by changes in phase synchronization in intracranial EEG recordings from epilepsy patients," *Epilepsia,* 2000, 41(7), pp. 167-172.