Computers in Biology and Medicine

Computers in Biology and Medicine 34 (2004) 561–575

http://www.intl.elsevierhealth.com/journals/cobm

Automatic spike detection in EEG by a two-stage procedure based on support vector machines

Nurettin Acır*, Cüneyt Güzeliş

Electrical and Electronics Engineering Department, Dokuz Eylül University, İzmir, Turkey

Received 20 May 2003; accepted 25 August 2003

Abstract

In this study, we introduce a two-stage procedure based on support vector machines for the automatic detection of epileptic spikes in a multi-channel electroencephalographic signal. In the first stage, a modified non-linear digital filter is used as a pre-classifier to classify the peaks into two subgroups: (i) spikes and spike like non-spikes (ii) trivial non-spikes. The pre-classification done in the first stage not only reduces the computation time but also increases the overall detection performance of the procedure. In the second stage, the peaks falling into the first group are aimed to be separated from each other by a support vector machine that would function as a post-classifier. Visual evaluation, by two experts, of 19 channel EEG records of 7 epileptic patients showed that the best performance is obtained providing 90.3% sensitivity, 88.1% selectivity and 9.5% false detection rate.

Keywords: EEG; Automatic spike detection; Autoregressive model; Digital filter; Support vector machines

1. Introduction

ElectroEncephaloGram (EEG) is a one dimensional and multi-channel signal obtained from the brain. EEG is widely used clinical tool in the diagnosing, monitoring and managing of neurological disorders related to epilepsy. Epilepsy may be defined as a symptom of paroxysmal and abnormal discharges in the brain that may be induced by a variety of pathological processes of genetic or acquired origin. This disorder is often characterized by sharp recurrent and transient disturbances of mental function and/or movements of different body parts that result from excessive discharges of groups of brain cells. The presence of epileptiform activity, which is distinct from the background EEG activity, confirms the diagnosis of epilepsy, although it can be confused with other disorders

* Corresponding author. Tel.: +90-232-4531008; fax: +90-232-4534279.
E-mail address: nurettin.acir@eee.deu.edu.tr (N. Acır).

0010-4825/$ - see front matter © 2003 Elsevier Ltd. All rights reserved.
doi:10.1016/j.compbiomed.2003.08.003

producing similar seizure-like activity. During seizures, the scalp EEG of patients with epilepsy is characterized by high-amplitude synchronized periodic EEG waveforms. In between seizures, epileptiform transient waveforms, which include spikes and sharp waves, are typically observed in the scalp EEG of such patients. An EEG spike, which is different from the background activity, has a pointed peak and duration of 20–70 ms [1]. Although it may occur alone, a spike is usually followed by a slow wave, which lasts 150–350 ms, forming what is known as spike and slow wave (SSW) complex [2].

The evaluation of EEG for the detection of epilepsy generally includes visual scanning of EEG recordings for these spikes and seizures by an experienced electroencephalographer (EEGer). This process is time consuming, especially in the case of long recordings. In addition, disagreement among readers of the same record is possible due to the subjective nature of the analysis [3]. The use of ambulatory monitoring, which produces 24 h or longer continuous EEG recordings, became more common, thus further increasing the need for an efficient automated detection method. Several attempts have been made to automate the spike and seizure detection process through computer based methods. In most of these detection systems, measurements of electrographic parameters of EEG waveforms, such as sharpness, slope, duration and amplitude are compared with thresholds representative of a typical true spike and seizure.

Various methods have been used for the detection of spikes, sharp waves and other transient signals in the scalp EEG. Mostly they are based on measuring the duration and sharpness of individual waves using their second derivatives [4–6]. Another spike detection system has been developed that is sensitive to states of EEG such as active wakefulness, quite wakefulness, desyncronized EEG, phasic EEG and slow EEG [7]. In addition to these methods, mimetic techniques have been widely used for detecting epileptiform discharges. For spike detection, mimetic methods usually decompose the EEG into waves or half waves by determining the extrema of the amplitudes. Each wave is then examined for its fit to a set of predetermined criteria, e.g. duration, amplitude, slope and sharpness [6,8]. Similarly, some filtering techniques have been proposed for spike detection [9]. All these studies have tried to find some standards for detecting spikes using objective criteria.

Artifacts should be eliminated for a robust detection of epileptiform spikes. With an EEG signal free of artifacts, a reasonably accurate detection of spikes and sharp waves is possible; however, difficulties arise with artifacts. This problem increases the number of false detections that commonly plague all automatic detection systems [10,11]. Many studies using artificial neural networks (ANN) approach for detecting EEG spikes have been reported [8,11–15]. ANN based spike detection systems basically use either of two different input representations: (1) the extracted EEG features or (2) the raw EEG signal. In the former case, the extracted spike features such as slope, duration, amplitude and sharpness are presented to the ANN for training and testing purposes [16]. The success of such a system depends on proper selection of the features, which is in some sense a trial and error procedure. In the second case, the raw EEG signal is presented to the ANN after a proper scaling and windowing [8,17,18]. Although it saves memory and time, the feature extraction based method is not very efficient in terms of the classification performance. On the other hand, the second method whose classification performance is superior to that of the first method requires a lot of memory and is very time consuming.

The novel support vector machine (SVM) based method proposed in this paper employs a two stage classification procedure that would display good classification performance while being efficient in terms of memory and time requirements. The former classification is realized by a modified non-linear digital filter which separates spikes and spike like non-spikes from trivial non-spikes. SVM has been examined as a post-classifier with radial basis kernel function [19,20]. SVM has a good performance resulting in 90.3% sensitivity, 88.1% selectivity and 9.5% false detection rate for the test set. In this study, the simulations were performed using MATLAB running on a Pentium Celeron 400 MHz PC computer.

2. Materials and methods

2.1. Recording and data acquisition

The EEG data used in this study were acquired from 25 epileptic patients who had been under evaluation and treatment in the Neurology Department of Dokuz Eylül University Hospital, İzmir, Turkey. 18 of these EEG records were used in training, while the remaining 7 in testing procedure. Data were obtained from a clinical EEG monitoring system, which stores continuous EEG data on its hard disk. EEG data were acquired with Ag/AgCl disk electrodes placed using the 10–20 international electrode placement system. The recordings were obtained from 19 channels with 256 Hz sampling frequency and band-pass filtered between 1 and 70 Hz. Data were then stored on both a hard disk and an optical disk.

2.2. Peak detection

Our approach is based on the fact that a spike reflects itself as a peak in the EEG record. Therefore, the first step would be the extraction of the peaks from the record. To detect peaks, we first took the average of the signal and subtracted it from the original signal. Then we calculated the time-derivative of the signal $f(t)$ and found its zero-crossings. This can be achieved for a discrete signal $f[n] = f(t)|_{t=nT}$ by searching inflection points based on the following criteria:

- If $f[n] - f[n-1] > 0 > f[n+1] - f[n]$, then this is a positive peak; we store it with index.
- If $f[n] - f[n-1] < 0 < f[n+1] - f[n]$, then this is a negative peak; we store it with index.
- If $[f[n] - f[n-1]].[f[n+1] - f[n]\ ] > 0$, then this is not a peak; we discard it.

The above procedure is indeed a numerical differentiation technique such that $(f(n+1) - f(n))/T$ represents first order forward approximation to the derivative of $f(t)$ at $t = nT$.

We used this procedure only for using in extracting features that would be fed to the post-classifier as explained in Section 3.

2.3. Pre-classification

Pre-classification procedure is performed to eliminate trivial non-spikes and also to capture spikes and spike like non-spikes. Spikes and spike like non-spikes are the only inputs of the post-classifier. In this way, not only the computation time of the entire classification procedure is reduced, but also the overall detection performance is increased.

2.3.1. *Nonlinear digital filter as the pre-classifier*

The EEG can be considered as a summation of stationary background waves (trivial non-spikes) and non-stationary spike waves (spikes and spike like non-spikes). The stationary waves can be expressed by an autoregressive (AR) model as follows:

$$x_n = \sum_{k=1}^{N} a_k x_{n-k} + v_n, \tag{1}$$

where x_n is a sampled stationary wave signal, a_k are the coefficients of AR model, N is the order of AR model and v_n is a white noise having zero mean and whose probability density function is assumed to be nearly Gaussian. On the other hand, the non-stationary spike waves are composed of large amplitude random impulsive signals which occur in low rate.

On the basis of these assumptions, non-stationary spike waves, spikes and spike like non-spikes, are detected by a non-linear digital filter shown in Fig. 1. Here, y_n represents the output of this filter, that are, the detected non-stationary spike waves. e_n is the prediction error for the original EEG signal which is obtained as $e_n = x_n - \hat{x}_n$ where $\hat{x}_n$ is the output of prediction filter which predicts the x_n by using Adaptive AR model [21]. $F(\cdot)$ is the normalization function which will be strictly positive after following procedure:

$$F(\cdot) = d_n = \left(\frac{e_n}{\sigma}\right)^2. \tag{2}$$

$G(\cdot)$ is a unipolar non-linear function which satisfies the following condition as shown in Fig. 2:

$$G(\cdot) \geqslant \varepsilon, \tag{3}$$

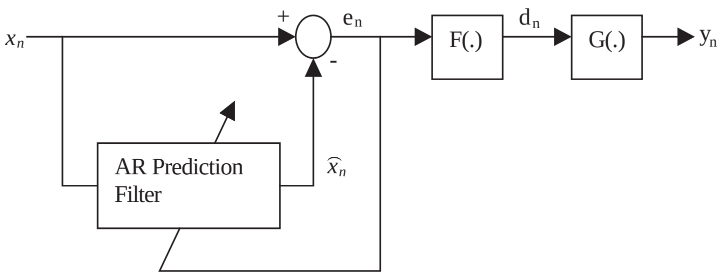

Fig. 1. Block diagram of a non-linear digital filter.

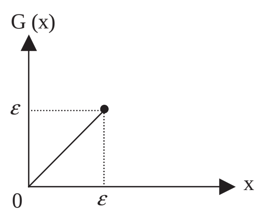

Fig. 2. Nonlinear function $G(\cdot)$.

where ε is a certain positive value which separates the non-stationary spike waves from background stationary waves which are trivial non-spikes. The ε value is determined by using the probability density function (PDF) of the normalized prediction error d_n. The probability of the detection of background stationary waves together with non-stationary spike waves was given as follows:

$$\int_{-\infty}^{\infty} p(x)\,\mathrm{d}x = 1, \tag{4}$$

where $p(x)$ is the PDF of the whole d_n. From (4), the probability of the detection of non-stationary spike waves (λ) can be expressed as in (5):

$$\lambda = 1 - \int_0^{\varepsilon} p(x)\,\mathrm{d}x, \tag{5}$$

where λ is the probability of the detection of non-stationary spike waves. Univariate normal density function is used as pdf which is completely specified by two parameters, the mean (μ) and variance (σ^2):

$$p(x) = \frac{1}{\sqrt{2\pi\sigma}}\,\mathrm{e}^{-(1/2)((x-\mu)/\sigma)^2} \tag{6}$$

for which

$$\mu = \int_{-\infty}^{\infty} x\,p(x)\,\mathrm{d}x \tag{7}$$

and

$$\sigma^2 = \int_{-\infty}^{\infty} (x-\mu)^2\,p(x)\,\mathrm{d}x. \tag{8}$$

Normally distributed samples tend to cluster about the mean, with a spread proportional to the standard deviation (σ); approximately 95% of the samples drawn from a normal population will fall in the interval $|x-\mu| \leqslant 2\sigma$ [22]. When we look at the distribution of the probability densities of d_n, the non-stationary spike waves cluster in far away from the mean with lower probability densities whereby stationary background waves cluster about the mean with higher probability densities. That mean non-stationary spike waves fall in the interval $|x-\mu| \geqslant 2\sigma$ which correspond to the about 5% of the samples which locate in a region having lower densities separated from the stationary background waves. So, ε value can be determined from (6) by corresponding to the $p(x)$ value from a certain upper bound of lower densities from about 5% λ region to $p(x)$. Thus, ε is determined adaptively which varies from data to data depending on the normalized prediction error distribution.

In this procedure, the input continuous EEG is firstly smoothed by using a low pass filter to eliminate the small peaks and high frequency noise [21]. Then, $\hat{x}_n$ is predicted on the basis of the AR model as follows:

$$\hat{x}_n = \sum_{k=1}^{N} a_k x_{n-k}. \tag{9}$$

After the prediction error for the input EEG is obtained ($e_n = x_n - \hat{x}_n$), e_n is normalized by using (2). Then, ε is determined as specified above. If $G(\cdot) \geqslant \varepsilon$, then x_n corresponding to that d_n is labelled as the spikes and spike like non-spikes (non-stationary spike waves), else it is labelled as trivial

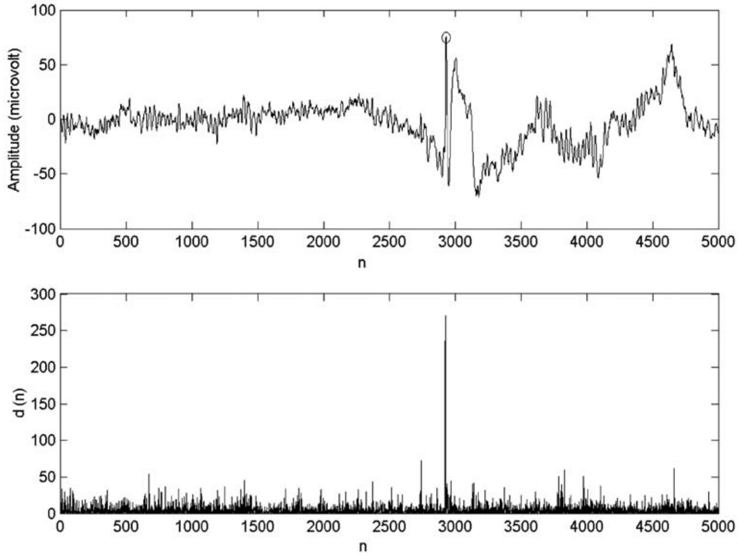

Fig. 3. An example of the detection of spikes in epileptic EEG by the non-linear digital filter.

non-spikes (background stationary waves). The order of the AR model (N) determined by using Akaike criterion [23] was set to 10 and $\lambda = 0.05$. The coefficients of the AR prediction filter are calculated by using least mean square algorithm [23].

As is shown in Fig. 3, the non-linear digital filter can detect the spikes in EEG effectively. However, some spike like non-spikes of the EEG other than the spikes are also detected in experiments. But, approximately all spikes can be detected by using non-linear digital filter by adjusting λ. Hence, this technique is used as an efficient preprocessing technique in the proposed automatic spike detection system.

By pre-classification, $\sim$ 90% of the peaks fall into the second group as trivial non-spikes. The remaining $\sim$10% of the peaks, which belongs to the first group, is given as input to the post-classifier. In this way, both the training set and training time of the post-classifier is drastically reduced.

2.4. Post-classification

The function of the post-classifier is to separate the peaks in the first group, i.e. spikes and spike like non-spikes, from each other. An SVM with a hard-limiter output unit is used as the post-classifier.

2.4.1. Support vector machines

Fig. 4 shows the architecture of the SVM. M is the number of support vectors.

SVM is a relatively new approach for solving supervised classification problems and is very useful due to its generalization ability. In essence, such an approach maximizes the margin between the training data and the decision boundary, which can be casted as a quadratic optimization problem. The subsets of patterns that are closest to the decision boundary are called support vectors.

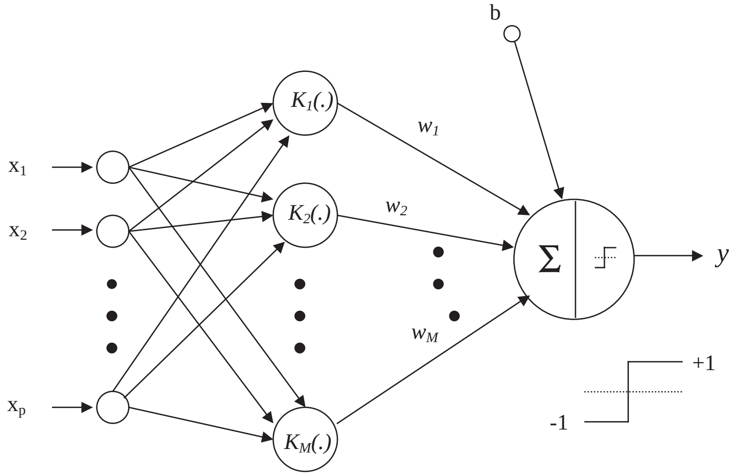

Fig. 4. Architecture of the SVM used for post-classification.

For a linearly separable binary classification problem, the construction of a hyperplane $\mathbf{w}^{\mathrm{T}}\mathbf{x}+b=0$ so that the margin between the hyperplane and the nearest point is maximized can be posed as the following quadratic optimization problem [24]:

$$\min_{\mathbf{w}} \frac{1}{2}(\mathbf{w}^{\mathrm{T}}\mathbf{w}) \tag{10}$$

subject to

$$d^s((\mathbf{w}^{\mathrm{T}}\mathbf{x}^s)+b) \geqslant 1 \quad \text{with } s=1,\ldots,S, \tag{11}$$

where $d^s \in \{-1,1\}$ stands for the sth desired output, $x^s \in R^p$ stands for the sth input sample of the training data set $\{x^s,d^s\}_{s=1}^{S}$. Eq. (10) forces a rescaling on $(\mathbf{w},b)$ so that the point closest to the hyperplane has a distance of $1/\|w\|$ [24]. Maximizing the margin corresponds to minimizing the Euclidean norm of the weight vector. Often in practice, a separating hyperplane does not exist. Hence constraint (11) is relaxed by introducing slack variables $\xi_s \geqslant 0$, $s=1,\ldots,S$. The optimization problem now becomes as follows (for a user defined positive finite constant C):

$$\min_{\mathbf{w},\xi} \frac{1}{2}(\mathbf{w}^{\mathrm{T}}\mathbf{w}) + C\sum_{s=1}^{S}\xi_s \tag{12}$$

subject to

$$d^s((\mathbf{w}^{\mathrm{T}}\mathbf{x}^s)+b) \geqslant 1-\xi_s, \tag{13}$$

$$\xi_s \geqslant 0 \quad \text{with } s=1,\ldots,S. \tag{14}$$

By introducing Lagrange multipliers α_S and using the Karush–Kuhn–Tucker theorem of optimization theory, we can pose the equivalent dual optimization problem [25]:

$$\max_{\alpha_s} \sum_{s=1}^{S}\alpha_s - \frac{1}{2}\sum_{r,s=1}^{S}\alpha_r\alpha_s d^r d^s((\mathbf{x}^r)^{\mathrm{T}}\mathbf{x}^s) \tag{15}$$

subject to

$$0 \leqslant \alpha_s \leqslant C \quad \text{with } s = 1, \ldots, S, \tag{16}$$

$$\sum_{s=1}^{S} \alpha_s d^s = 0. \tag{17}$$

The solution is given by

$$\mathbf{w} = \sum_{s=1}^{S} d^s a_s \mathbf{x}^s. \tag{18}$$

The non-zero α_s's correspond to the so-called support vectors $\mathbf{x}^s$ that help to define the boundary between the two classes. All other training examples with corresponding zero α_s values are now rendered irrelevant and automatically satisfy constraint (13) with $\xi_s = 0$. The hyperplane decision function can be written, for the vector $\mathbf{x}$, as follows:

$$f(\mathbf{x}) = \operatorname{sgn}\left(\sum_{s=1}^{S} d^s \alpha_s (\mathbf{x}^{\mathrm{T}} \mathbf{x}^s) + b\right). \tag{19}$$

To allow for more general decision surfaces, the inner product $\langle \mathbf{x}, \mathbf{x}^s \rangle = \mathbf{x}^{\mathrm{T}} \mathbf{x}^s$ can simply be replaced by a suitable kernel function $K(\cdot,\cdot)$. Hence the objective function to be maximized can now be written as:

$$\sum_{s=1}^{S} \alpha_s - \frac{1}{2} \sum_{r,s=1}^{S} \alpha_r \alpha_s d^r d^s K(\mathbf{x}^r, \mathbf{x}^s) \tag{20}$$

with the constraint equations (16) and (17) unchanged. The decision function, for the vector $\mathbf{x}$, then becomes:

$$f(\mathbf{x}) = \operatorname{sgn}\left(\sum_{s=1}^{S} d^s \alpha_s K(\mathbf{x}, \mathbf{x}^s) + b\right). \tag{21}$$

The α_ss are determined from (20), (16), and (17). The bias parameter b is determined from (19) by using two arbitrary support vectors from known but opposite classes.

By replacing the inner products with kernel functions, the input data are mapped to a higher dimensional space. It is then in this higher dimensional space that a separating hyperplane is constructed to maximize the margin. In the lower dimensional data space, this hyperplane becomes a non-linear separating function.

3. Results and discussion

To visualise the problem and to see the effect of C parameter, we restrict ourselves to the two features that contain the most important information about the class, namely the duration and the

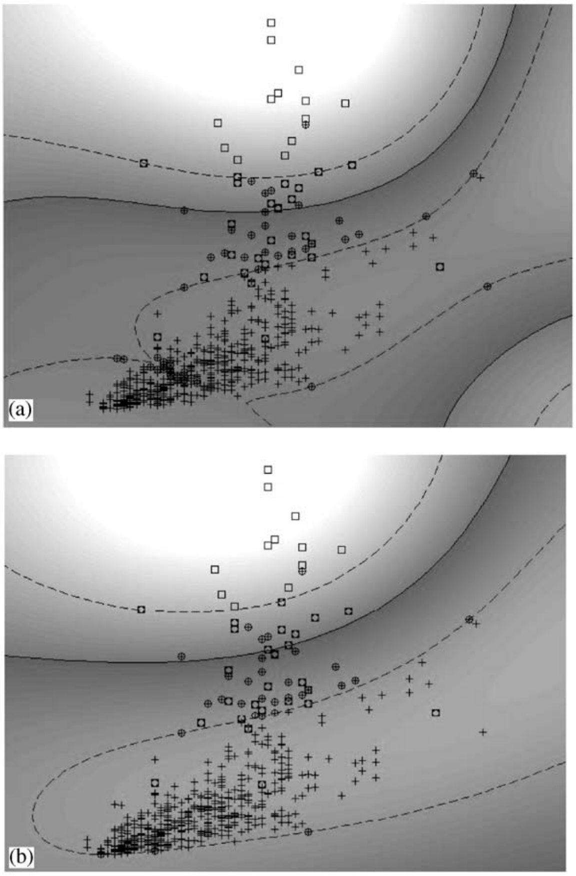

Fig. 5. An example of the effect of C on the separation of spikes: solid line represents decision boundary, dashed line represents maximized margins, (□) represents spike and (+) represents non-spike: (a) $C = 100$; (b) $C = 0.1$.

amplitude. As can be seen from Fig. 5, the separation of a group of spikes from a group of non-spikes is not so trivial. In the SVM classification, support vectors are represented by circles as shown in Fig. 5. Fig. 5 shows the results of an SVM classification for two different degrees of misclassification tolerance, so it visualises the effect of the tolerance on misclassification errors on the topology of the classifier boundary. After some trials, the value of $C = 100$ gives the best result. In this implementation, we construct an SVM by using radial basis function as the kernel function in SVM:

$$K(\mathbf{x}, \mathbf{x}^s) = \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}^s\|}{2\sigma^2}\right), \tag{22}$$

where $\sigma^2 = 1$.

After the pre-classification stage, a new data set is constructed by extracting new features of the spike candidates from the original signal by a procedure as will be explained below. In this

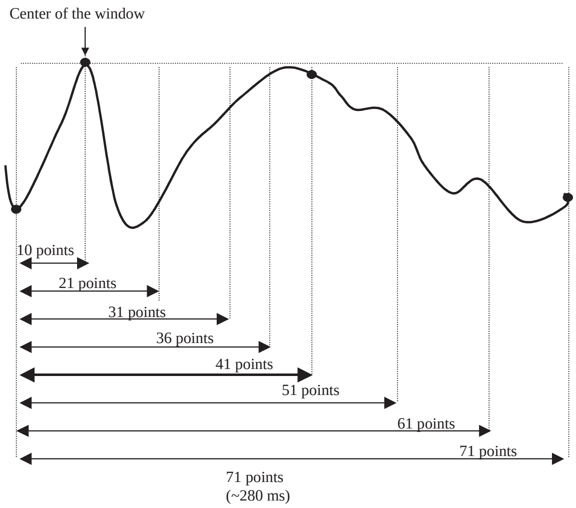

Fig. 6. A sketch of a spike extracted from the original signal and used as the input to the SVM using different number of data points from the right of the center.

procedure, the time indices of data have been stored and peaks have been determined as explained in Section 2.2. The procedure is as follows:

- The main peaks of the spike candidates are taken as the center of the window.
- 10 data points from left and 10 data points from right of the center are recalled by using the previously stored time indices from the original data.
- This procedure is repeated for 20, 30, 35, 40, 50, 60 and 70 data points from the right of the center, while 10 data points for left is fixed (Fig. 6).

Using the above procedure, the training set is prepared without any modification on spike forms by extracting the spikes from the original signal in the determined intervals.

The classification performance of the system is determined measuring the sensitivity and selectivity of the overall classifier. Sensitivity is the ratio of true positives to the total number of true spikes detected by EEGer. Selectivity is the ratio of true positives to the total number of spikes detected by our system. Spikes are called as "true positive" when both our system and experts detect them as spike [26].

To see the effect of input window size on the performance of SVM, the implementation is repeated for different input window sizes (Table 1). The results presented in Table 1 imply that, the input size of 41 data points which correspond to time duration of 160 ms is the optimal choice. Spike morphology studies also show that the spike duration is, on the average, about 75 ms [27]. Positioning of such a spike wave form in a 160 ms window provides all the information to the SVM including the wave components.

It is observed that the window size which is equal to the input dimension of the SVM affects the performance of the SVM as the post-classifier.

Table 1
The effects of input size on the overall classification performance

Input size	Sensitivity (∼ %)	Selectivity (∼ %)
21	79.7	74.6
31	84.5	78.5
36	87.6	85.8
41	90.3	88.1
51	90.3	88.0
61	90.2	87.9
71	90.1	87.6

Table 2
Test results of the SVM for 7 epileptic patients

Patient	Age	Duration	No. of spikes in test set	SVM Sensitivity (∼ %)	SVM Selectivity (∼ %)
1	14	31.6 min	9	88.8	88.8
2	18	69.0 min	41	87.8	83.7
3	52	25.3 min	3	100.0	100.0
4	16	44.3 min	16	87.5	87.5
5	22	38.0 min	22	90.1	87.0
6	28	34.8 min	16	93.8	88.2
7	13	50.6 min	32	84.4	81.8
Total	∼ 23	∼ 4 h 14 min	139	∼ 90.3	∼ 88.1

The data set contains 41 data points for each peak that is selected after the pre-classification stage. Spike and non-spike events are represented by $+1$ and -1, respectively, for both training and testing procedures. We have trained the machine for different C values until to have the best result. The best result is obtained for $C = 100$ in the testing procedure. The number of support vectors is 29 which corresponds to 3% of the training data. 7 epileptic records are tested. Classification performance of the SVM is calculated by measuring its sensitivity and selectivity. Testing the SVM shows a sensitivity of 90.3% and a selectivity of 88.1% (Table 2).

We also calculate the false detection rate, which is a method used in the diagnosis of neurological disorders related to epilepsy, to determine the performance of the systems (Table 3). The false detection rate can be defined as the ratio of the false detected peaks to the total number of spikes detected by the experts. As can be seen from Table 3, the false detection rate of SVM procedure is low. This implies that, SVM can be used effectively in the accurate diagnosis of epilepsy.

Fig. 7 shows an example of spikes detected by the SVM in result of the overall detection procedure.

Table 3
The false detection rates of the SVM

Patient	Age	Duration	SVM (∼ %)
1	14	31.6 min	11.1
2	18	69.0 min	14.1
3	52	25.3 min	0.0
4	16	44.3 min	12.5
5	22	38.0 min	9.0
6	28	34.8 min	8.5
7	13	50.6 min	11.6
Total	∼ 23	∼ 4 h 14 min	∼ 9.5

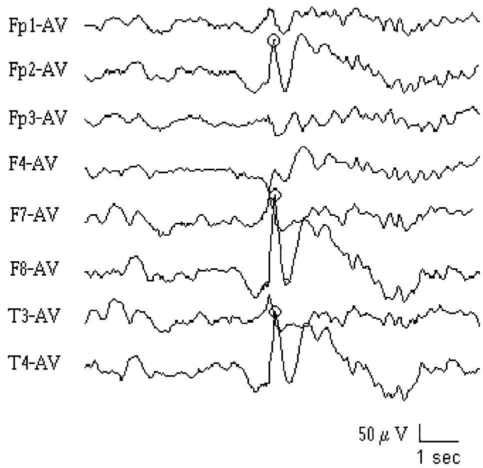

Fig. 7. An example of spikes detected by the post-classification procedure.

4. Conclusion

In this study, we introduce a novel two stage classification procedure based on SVM for spike detection that will contribute to the clinical applications with its good efficiency and accuracy levels. The proposed approach accomplishes peak detection, feature extraction, pre-classification, and spike detection by preserving the original form of the spikes. We demonstrate that the pre-classification done by the non-linear digital filter can successfully separate the spikes and spike like non-spikes from trivial non-spikes. The pre-classification not only reduces the computation time but also increases the overall detection performance of the SVM. Using the output of the pre-classification stage as the input to the post-classification procedure, the best performance is obtained with SVM providing 90.3% sensitivity, 88.1% selectivity and 9.5% false detection rate. SVM in two stage classification procedure achieves a significant improvement in terms of sensitivity, selectivity and false detection rate.

5. Summary

In this study, we introduce a two-stage procedure based on support vector machines for the automatic detection of epileptic spikes in a multi-channel electroencephalographic signal. In the first stage, a modified non-linear digital filter is used as a pre-classifier to classify the peaks into two subgroups: (i) spikes and spike like non-spikes (ii) trivial non-spikes.

Pre-classification procedure is performed to eliminate trivial non-spikes and also to capture spikes and spike like non-spikes. Spikes and spike like non-spikes are the only inputs of the post-classifier. In this way, not only the computation time of the entire classification procedure is reduced, but also the overall detection performance is increased. Pre-classification procedure is composed of the following steps. By pre-classification, $\sim$ 90% of the peaks fall into the second group as trivial non-spikes. The remaining $\sim$10% of the peaks, which belongs to the first group, is given as input to the post-classifier. In this way, both the training set and training time of the post-classifier is drastically reduced.

In the second stage, the peaks falling into the first group are aimed to be separated from each other by a support vector machine that would function as a post-classifier. Support vector machine (SVM) is a relatively new approach for solving supervised classification problems and is very useful due to its generalization ability. In essence, such an approach maximizes the margin between the training data and the decision boundary, which can be casted as a quadratic optimization problem. The subsets of patterns that are closest to the decision boundary are called support vectors.

The proposed approach accomplishes peak detection, feature extraction, pre-classification, and spike detection by preserving the original form of the spikes. We demonstrate that the pre-classification unit trained by the non-linear digital filter can successfully separate the spikes and spike like non-spikes from trivial non-spikes. Using the output of the pre-classification stage as the input to the post-classification procedure, the best performance is obtained with SVM providing 90.3% sensitivity, 88.1% selectivity and 9.5% false detection rate. SVM in two stage classification procedure achieves a significant improvement in terms of sensitivity, selectivity and false detection rate.

Acknowledgements

The authors gratefully acknowledge the help of Neurology Department of the Dokuz Eylül University in the preparation of the database. The first author Nurettin Acır would also like to thank to TÜBİTAK (Turkish Scientific and Technical Research Council) Münir Birsel Fund for support.

References

[1] E. Chatrian, L. Bergamini, M. Dondey, D.W. Klass, M. Lennox-Buchthal, I. Petersen, A glossary of terms most commonly used by clinical electroencephalographs, Electroencephalogr. Clin. Neurophysiol. 37 (1974) 538–548.

[2] T. Kalaycı, Ö. Özdamar, Wavelet preprocessing for automated neural network detection of EEG spikes, IEEE Eng. Med. Biol. 14 (2) (1995) 160–166.

[3] P.Y. Ktonas, Methods of Analysis of Brain Electrical and Magnetic Signals, EEG Handbook, Elsevier, Amsterdam, 1987.

[4] D.O. Walter, H.F. Muller, R.M. Jell, Semiautomatic quantification of sharpness of EEG phenomenon, IEEE Trans. Biomed. Eng. 20 (1973) 53–54.
[5] P.Y. Ktonas, J.R. Smith, Quantification of abnormal EEG spike characteristics, Comput. Biol. Med. 4 (1974) 157–163.
[6] J. Gotman, P. Gloor, Automatic recognition and quantification of interictal epileptic activity in the human scalp EEG, Electroencephalogr. Clin. Neurophysiol. 41 (1976) 513–529.
[7] J. Gotman, L.Y. Wang, State-dependent spike detection: concepts and preliminary results, Electroencephalogr. Clin. Neurophysiol. 79 (1991) 11–19.
[8] A.A. Dingle, R.D. Jones, G.J. Carroll, W.R. Fright, A multistage system to detect epileptiform activity in the EEG, IEEE Trans. Biomed. Eng. 40 (1993) 1260–1268.
[9] G. Pfurtscheller, G. Fischer, A new approach to spike detection using a combination of inverse and matched filter techniques, Electroencephalogr. Clin. Neurophysiol. 44 (1977) 243–247.
[10] D. Frost, Automatic recognition and characterization of epileptiform discharges in human EEG, J. Clin. Neurophysiol. 2 (3) (1985) 231–249.
[11] R. Glover, N. Raghavan, P.Y. Ktonas, J.D. Frost, Context-based automated detection of epileptogenic sharp transients in the EEG: elimination of false positives, IEEE Trans. Biomed. Eng. 36 (5) (1989) 519–527.
[12] A.J. Gabor, M. Seyal, Automated interictal EEG spike detection using artificial neural networks, Electroencephalogr. Clin. Neurophysiol. 83 (1992) 271–280.
[13] C.W. Ko, Y.D. Lin, H.W. Chang, G.J. Jan, An EEG spike detection algorithm using ANN with multi-channel correlation, Proceedings of the 20th Annual International Conference Of the IEEE Engineering in Med and Biol Society, Vol. 20(4) (1998), pp. 2070–2073.
[14] T. Shimada, T. Shiina, Y. Saito, Detection of characteristic waves of sleep EEG by neural network analysis, IEEE Trans. Biomed. Eng. 47 (3) (2000) 369–379.
[15] C.W. Ko, H.W. Chung, Automatic spike detection via an artificial neural network using raw EEG data: effect of data preparation and implications in the limitations of online recognition, Clin. Neurophysiol. 111 (2000) 477–481.
[16] G. Hellmann, Multifold features determine linear equation for automatic spike detection applying neural nin interictal EcoG, Clin. Neurophysiol. 110 (1999) 887–894.
[17] F.M.C. Besag, M. Mills, F. Wardale, C.M. Andrew, M.D. Craggs, The validation of a new ambulatory spike and wave monitor, Electroencephalogr. Clin. Neurophysiol. 73 (1989) 157–164.
[18] Ö. Özdamar, T. Kalayci, Detection of spikes with artificial neural networks using raw EEG, Comput. Biomed. Res. 31 (1998) 122–142.
[19] V. Vapnik, S. Golowich, A. Smola, Support vector method for function approximation, regression estimation, and signal processing, Advances in Neural Information Processing Systems, 9, MIT Press, Cambridge, MA (1997), pp. 281–287.
[20] B. Boser, I. Guyon, V. Vapnik, A training algorithm for optimal margin classifiers, Fifth Annual Workshop on Computational Learning Theory, ACM Press, New York, 1992, pp. 144–152.
[21] M. Akay, Biomedical Signal Processing, Academic Press, San Diego, 1994.
[22] R.O. Duda, P.E. Hart, Pattern Classification and Scene Analysis, Wiley, New York, 1973.
[24] S. Haykin, Neural Networks; A Comprehensive Foundation, New Prentice-Hall, Englewood Cliffs, NJ, 1999.
[23] S. Haykin, Adaptive Filter Theory, Prentice-Hall, Englewood Cliffs, NJ, 1996.
[27] W.T. Blum, J.F. Lemieux, Morphology of spikes in spike and wave complexes, Electroencephalogr. Clin. Neurophysiol. 69 (1988) 508–513.
[25] D.P. Bertsekas, Nonlinear Programming, Athenas Scientific, MA, Belmont, 1995.
[26] W.R.S. Webber, B. Litt, R.P. Lesser, R.S. Fisher, I. Bankman, Automatic EEG spike detection: what should the computer imitate, Electroencephalogr. Clin. Neurophysiol. 87 (1993) 364–373.

Nurettin Acır graduated from Erciyes University, Turkey, in 1995. He took M.Sc. degree from Nigde University, Turkey, in 1998, all in Electronics Engineering. He is going on Ph.D. studies at Electrical and Electronics Engineering Department of Dokuz Eylül University, Turkey. He is now a visiting scholar at Neurosensory Engineering Lab., University of Miami, USA. His interest areas include intelligent systems, biomedical signal processing, artificial neural networks.

Cüneyt Güzeliş graduated from İstanbul Technical University, Turkey, in 1981. He took M.Sc. and Ph.D. degree from İstanbul Technical University, Turkey, in 1984 and 1988, respectively, in all electrical engineering. Between 1989 and 1991 he worked in the Department of Electrical and Computer Engineering at University of California, USA, as visiting researcher and lecturer. He is now a professor at Electrical and Electronics Engineering Department of Dokuz Eylül University, Turkey. His interest areas include neural networks, signal processing and nonlinear circuits and systems.