ISSN: 0975-8585

Research Journal of Pharmaceutical, Biological and Chemical Sciences

EEG Signal Classification and Segmentation for Automated Epileptic Seizure Detection using SVM Classifier

Suguna Nanthini B*, and B Santhi.

School of Computing, SASTRA University, Thanjavur, Tamil Nadu 613401, India.

ABSTRACT

Epilepsy is an unusual brain activity due to abnormal behavior of neuronal discharge which occurs from concurrent seizures. The electroencephalogram (EEG) is a test that measures the electrical movement in the brain. The main objective of this EEG analysis is to introduce some novel methods for seizure detection and to compare the performance of the Support Vector Machine (SVM) classifier kernels. The 8 statistical features namely mean, standard deviation, median, mode, skewness, kurtosis, maximum and minimum and the 4 Gray Level Co-occurrence Matrix (GLCM) features namely contrast, correlation, energy and homogeneity are extracted from each segment of the EEG signal which are analyzed and classified using two different SVM kernels. The Linear and Radial Basis Function (RBF) kernels are used with the SVM classifier. The performance of the kernels is compared and concluded that the linear kernel is the best choice for this study of seizure detection. Further in this analysis, the parameter sigma in RBF kernel is tuned to the value two and compared with the default sigma value one. It seems that the accuracy is better in classification when the sigma value is tuned.

Keywords: EEG, Seizure, Segment, SVM, Linear, RBF

**Corresponding author*

INTRODUCTION

Brain is the essential part which controls the overall nervous system in human body. The electricity produced by the neurons in the brain is measured by EEG test. The abnormal electrical activity is the symptom of the brain disorder [16] [18]. Epilepsy also referred to as seizure disorder, because of the repeated and random interruptions of the normal brain functions and it can be identified by EEG [1]. EEG is a useful test for measuring the electrical activity and diagnosing epilepsy [13] [15].The number of processing modules was included in this study as follows; preprocessing EEG signal; dividing the EEG signal into four segments with equal in length; extraction of 12 dimensional features (The 8 statistical features namely mean, standard deviation, median, mode, skewness, kurtosis, maximum and minimum and the 4 Gray Level Co-occurrence Matrix (GLCM) features namely contrast, correlation, energy and homogeneity) and classification using SVM classifier (Figure 1). To add a novelty, this study concentrated by dividing the EEG data into different segments and extracted the features from each segment for detecting the seizure.

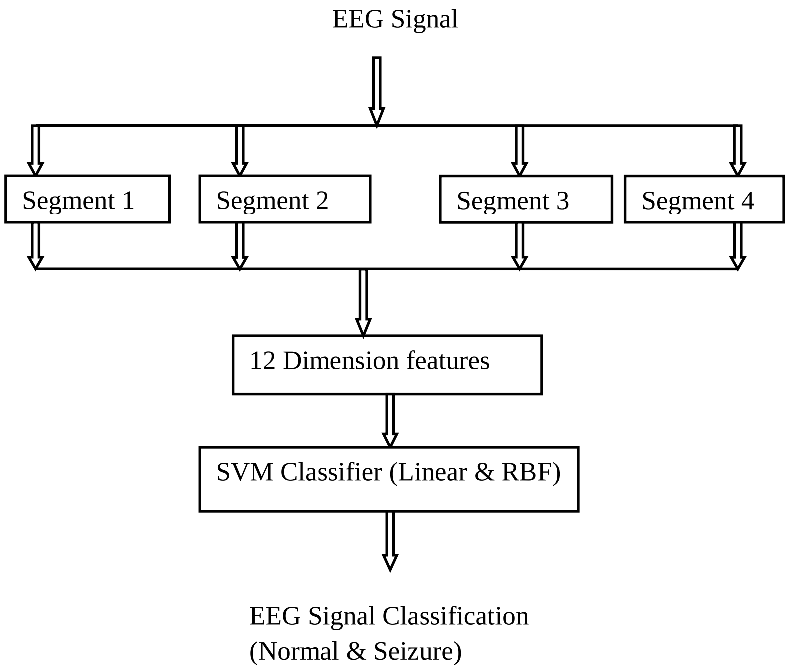


Figure 1: Processing modules for EEG signal classification

Kernel Ridge Regression (K-RR), Kernel Partial Least Squares (K-PLS), Kernel Principal Component Regression (K-PCR) and Support Vector Machines (SVM) are the popular techniques which can apply a kernel function for machine learning and data mining fields [29]. In this analysis, the EEG signals were divided into 4 segments with equal length. The 8 statistical features and the 4 GLCM features were extracted from each segment of the EEG signal. The EEG signals were analyzed and classified using two different SVM kernels.

MATERIALS AND METHODS

EEG Database

The Bonn University, Germany database [2] is downloaded for the purpose of EEG signal analysis. The information on this database is given in the Table 1. The segments of these 5 datasets A to E, each contains 100 single-channel EEG signals of 23.6s duration measured by the standardized electrode placement scheme. These segments were chosen under visual inspection. Using an average common reference with 128-channel amplifier most of the EEG signals were measured. The data dimensions were digitized using 12 bit A/D resolution at 173.61 Hz sampling rate. Datasets A, B and E were taken in this analysis for the binary classification. The sample from a single subject of EEG signals from each set (A-E) is shown in the following Figure 2.

Feature Extraction

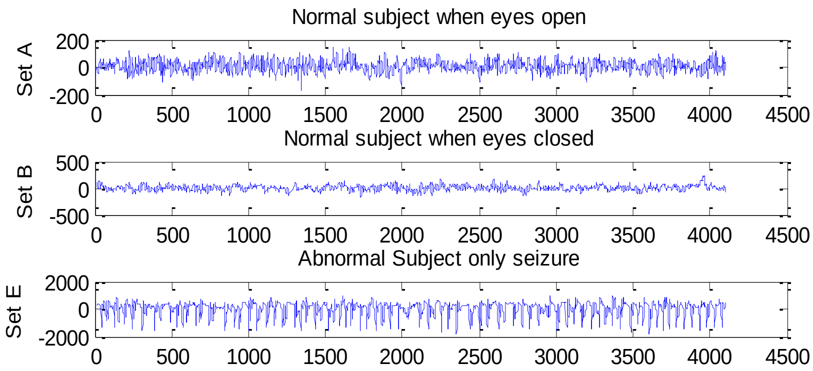
GLCM features

To reduce the dimension of EEG data, feature extraction is a unique form has been used in machine learning for the classification. In image processing, GLCM features are used for texture analysis. Nanthini and Santhi, [10] [11] indicates significant of GLCM features in EEG signal analysis. EEG signals are used to construct GLCM matrix.

Table 1: Information from Bonn University Database

Datasets	Description
Set A	EEG data recorded using a 10-20 system when the healthy persons were resting in the eyes open relaxation state with surface electrode type.
Set B	Using the same circumstance, EEG data recorded when the healthy persons were resting in the eyes closed relaxation state with surface electrode type.
Set C	EEG data recorded opposite to the epileptogenic zone of electrode placement when the persons were in Interictal state (seizure free intervals).
Set D	EEG data recorded within the epileptogenic zone of electrode placement when the persons were in Interictal state (seizure free intervals).
Set E	EEG data recorded within the epileptogenic zone of electrode placement when the persons were in Ictal state (merely seizure state).

Figure 2: Sample EEG Datasets



The texture characteristics (contrast, correlation, energy and homogeneity) are extracted and used for the classification. These texture representations which can applicable to signal are selected for the feature extraction process. The number of gray levels G presented in EEG signal is equal to the number of rows and columns in GLCM matrix. Therefore matrix element in the signal $S\ (i, j \mid \Delta x, \Delta y)$ is the two pixels of the relative frequency divided by a pixel distance $(\Delta x, \Delta y)$ where i and j are intensities occur within a given neighbourhood. Considering for the given n*m neighbourhood of an input containing *G* gray levels from *0* to *G-1*, let $I\ (p, q)$ be the intensity at sample m, line n of the neighbourhood. The co-occurrence matrix *C* is defined over $n * m$ and *I* of EEG signal is parameterised by an offset $(\Delta x, \Delta y)$ as follows:

$$C_{\Delta x \Delta y}(i,j) = \sum_{p=1}^{n} \sum_{q=1}^{m} I \begin{cases} 1, if I(p,q)=i and I(p+\Delta x, q+\Delta y)=j \\ 0\ otherwise \end{cases} \qquad (1)$$

where *i* and *j* are the EEG intensity values of the signal, *p* and *q* are the spatial positions in the EEG *I*. The offset$(\Delta x, \Delta y)$ is a reference pixel and its immediate neighbours depend on the direction used. The GLCM is a two dimensional array which obtains the specific position of a pixel compared to other pixels [12]. In this analysis, four GLCM features are used for the classification and they have been explained as follows.

In the measure of contrast, the local intensity varies in $S\ (i, j)$ where $i \neq j$, therefore this occurs away from the diagonal and defined as:

$$CONTRAST = \sum_{n=0}^{G-1} n^2 \left\{\sum_{i=1}^{G} \sum_{j=1}^{G} S\ (i,j)\right\} \quad (2)$$

Correlation is defined as a measure of gray scale. It is linearly dependent between the pixels at specified spots qualified to each other. Here in signal analysis, it is an important feature to see how the EEG vectors are fluctuating in the signal wave:

$$CORRELATION = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \frac{\{i*j\}*S(i,j)-\{\mu_x*\mu_y\}}{\sigma_x*\sigma_y} \quad (3)$$

Energy takes the smallest value when all the entries are equal. It is also called as uniformity and defined as follows:

$$ENERGY = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} S\ (i,j)^2 \quad (4)$$

Angular Second Moment (ASM) is a homogeneity that measures the signal$S\ (i,j)$. Taking less gray levels from homogeneous scene, GLCM gives higher values of $S(i,j)$. High values are obtained by using the following formula:

$$ASM\ (Homogeneity) = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \{S(i,j)\}^2 \quad (5)$$

Therefore, the values from the Equation 2, 3, 4 and 5 are applied as the input to the classifier.

Statistical features

As done in our previous study [37], the same available online Bonn University database [4] had been analyzed for this study. Each segment of EEG signal is analyzed with the above said 12 dimensional features. These features act as input to the SVM classifier for signal classification. The short description of each statistical feature is explained as follows:

MEAN: The value that represents the average of the EEG vectors and it is calculated as:

$$\bar{S} = \frac{\sum_{i=1}^{n} S_i}{n} \quad (6)$$

Here S_i denotes the EEG vectors.

MEDIAN: The mean of the two middle values from the sample EEG vectors (In the database, each sample of EEG signal contained 4096 vectors).

MODE: The point which appears most frequently in the set of EEG vectors.

STANDARD DEVIATION: To quantify the amount of variation, the value σ has been measured from EEG vectors.

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(S_i-\bar{S})^2}{n}} \quad (7)$$

SKEWNESS: This represents the measure of symmetry.

$$Skew(S) = \frac{\Sigma\left(\frac{Si-\bar{S}}{\sigma}\right)^3}{n} \quad (8)$$

KURTOSIS: To measure the tailedness, the kurtosis is calculated using the following formula.

$$Kur(S) = \frac{\Sigma\left(\frac{Si-\bar{S}}{\sigma}\right)^4}{n} - 3 \quad (9)$$

MAXIMUM: The maximum value from the EEG vectors

MINIMUM: The minimum value from the EEG vectors

The extracted statistical and GLCM features have been combined and given to SVM classifier. Then this system is trained well for the signal classification. The k fold cross validation is the process for segregating the input data into k subsets. Among the k subsets, k-1 subsets are utilized to train the system and the remaining one subset is used for the testing stage. This procedure will be repeated for k times (folds) [5] [17]. After all subsets completely validated, the results achieved from the k folds could be averaged to conclude the performance accuracy. The features are separated into two different classes with respect to their features label [19]. In this analysis, the samples from set A with E and set B with E for binary classification have been validated by using the 10-fold cross validation method. Then classification technique is processed for seizure detection. In machine learning, the kernel trick is commonly used for transforming input field to feature field. The linear methods are exercised to find out the optimum result of the classification process. The performance of the SVM kernels namely linear and RBF is compared in this analysis for EEG signal classification.

SVM classifier on EEG signal

The extracted GLCM features from the EEG signals are inputted to the SVM classifier. SVM algorithm [3] creates the hyper plane to classify the normal and seizure features. It can deal with high dimensional data. Basically SVM is a linear classifier which classifies the two different classes efficiently. The features of the two classes are categorized by the labels “-1” and “+1” (figure 3). The features which are extracted from the signal is defined as,

$$S = \{(x_i, y_i)_{i=1}^{n}\} \quad (10)$$

where y_i is the label related to the pattern x_i and n stands for the number of samples.

Dot product or the scalar product of linear classifier is defined as,

$$W^T(x) = \sum_i w_i \, x_i \quad (11)$$

This Equation 11 can be considered in the function form as,

$$f(x) = W^T(x) + b \quad (12)$$

where w_i stands for weight vector and b stands for the bias. For case b=0, the set of vectors in $W^T(x) = 0$ produce hyperplane through the origin which divides the features into two classes. Kernel is an algorithm which can produce non-linear decision boundaries [6] [12]. Replacing the normal SVM (linear kernel) dot product with a kernel function is defined as Gaussian radial basis function classifier which is expressed as,

$$k(x_i x_j) = e^{-\|x_i - x_j\|^2 / 2\sigma^2} \quad (13)$$

The variables x_i and x_j represent the two sample data from the dataset. The default sigma value is one which has been associated to all the attributes in the dataset. The features are separated into two different classes with respect to their features label [19].

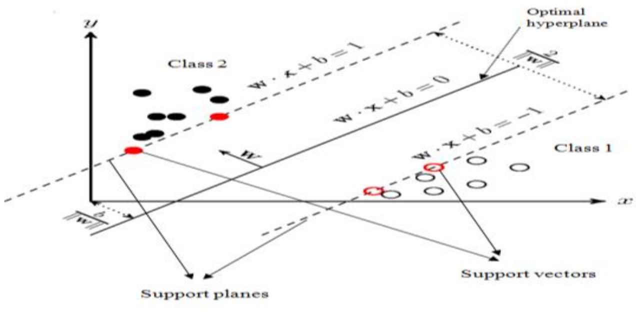


Figure 3: SVM classifier for seizure detection

Performance Evaluation

For this experiment a value "-1" indicates normal EEG pattern and "+1" points out the seizure EEG types (epileptic seizure). Most of the time, the performance is defined in terms of accuracy. Even though various methods are available to measure the performance, perfect measure does not exist. The performance evaluation of the ANN [14] classifier is examined by using the method called confusion matrix. While analyzing the output data, it is convenient to define three parameters (1) sensitivity is the True Positive Ratio (TPR) which identifies the truthfulness, (2) specificity is the True Negative Ratio (TNR) which identifies the misclassification and (3) total classification accuracy which defines the final performance measure of the classifier. They are numerically computed using the following formulae.

$$\text{Sensitivity} = \text{TPR} = \frac{TP}{TP+FN} * 100\% \quad (14)$$

$$\text{Specificity} = \text{TNR} = \frac{TN}{TN+FP} * 100\% \quad (15)$$

$$\text{Accuracy} = \frac{(TP+TN)}{(TP+FP+TN+FN)} * 100\% \quad (16)$$

where TP, TN, FP and FN are referred for true positive, true negative, false positive and false negative in the order mentioned.

RESULTS AND DISCUSSION

The pseudo code for this analysis is as follows.

Step 1: Receiving EEG signal from the source.

Step 2: After simple pre-processing, the EEG signal is divided into four segments which are equal length.

Step 3: The above said 12 dimension features are extracted from each segment.

Step 4: The system is trained in every combination of segment from the normal set and seizure set for binary classification.

Step 5: The performance accuracy of the SVM classifier with two different kernels namely Linear and RBF is compared for the EEG signal classification.

Step 6: Further the parameter sigma in Equation 13 (RBF kernel) is tuned to the value two and is compared with the default value one

Step 7: The analysis proves that the RBF kernel gets good result for EEG signal classification by tuning the sigma value.

Step 8: Finally, the study declares that the linear kernel is the best choice for this analysis.

Table 2 shows the sample features from normal and seizure subjects. Table 3 explains the performance of the SVM classifier with different kernels. The extracted maximum values from normal and seizure samples are plotted and shown in the figure 4. Figure 5 shows the binary classification using the segments of datasets A, B and E. From the result of Table 3, the classifier shows very good accuracy while using set A with set E. Set A includes normal EEG recordings when the person's eyes are open on relaxed state. Normally the EEG signals are monitored in two kinds of resting stages (eyes are opened and eyes are closed). The clinical reports confirm dramatic effects on these EEG recordings [9]. The purpose of this analysis has to be verified if there is any change in the performance of the classifier while using these sets of EEG recordings. But this analysis achieves no drastic changes in the performance of the classifier (Table 3). The study yields consistent accuracy values in every combination of segment. The features of normal and seizure EEG signals are linearly well separated because the input EEG vectors are linear. Thus the study concludes that the SVM classifier with linear kernel is the best choice for this EEG analysis. The signals from normal and seizure persons during EEG test are exactly describing their nature and perfectly coincide with their features. Therefore the

ISSN: 0975-8585

classification accuracy in linear kernel is in successful state. The next part of the analysis will focus on RBF kernel. This part of the study confirms when the parameter sigma in the Equation 13 is tuned to a value 2 the system's performance seems to be better than in the default sigma value 1. The EEG signal is directly manipulated in this study.

Table 2: Sample values of statistical features.

Features	Subjects			
	Normal1	Normal2	Seizure1	Seizure2
Mean	7.42846	53.8673	44.0487	29.6078
Median	6.34421	51.4898	61.7965	6.61957
Mode	-38.531	151.306	-817.46	-852.598
Standard Deviation	19.7646	33.4459	277.092	354.774
Skewness	0.22147	0.36673	0.50919	0.10145
Kurtosis	2.88368	2.72679	2.75450	2.47013
Maximum	59.5775	8.56668	579.069	895.605
Minimum	-38.531	151.306	-817.46	852.598
Contrast	18.4153	2.50201	22.4606	17.9465
Correlation	0.18245	0.05366	0.05243	0.26704
Energy	0.28183	0.87802	0.26278	0.26756
Homogeneity	0.65542	0.94681	0.59540	0.67900

Table 3: Performance measures of SVM kernels

Dataset	Kernels		
	Linear (%)	RBF (%)	Tuned RBF (%)
A with E	99.95	98.6	99.59
B with E	99.25	98.69	99.41

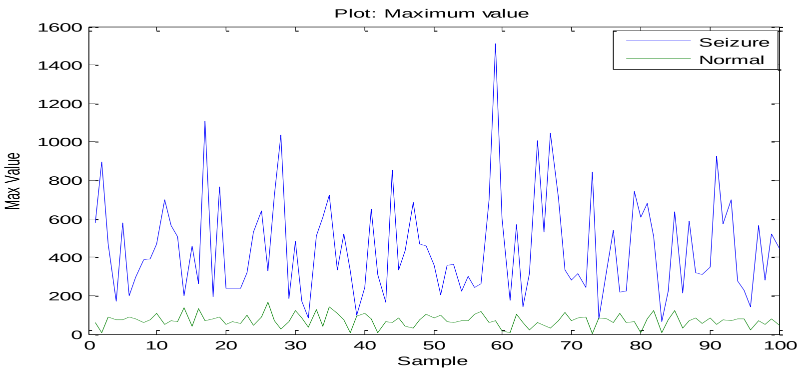


Figure 4: Max values from EEG samples

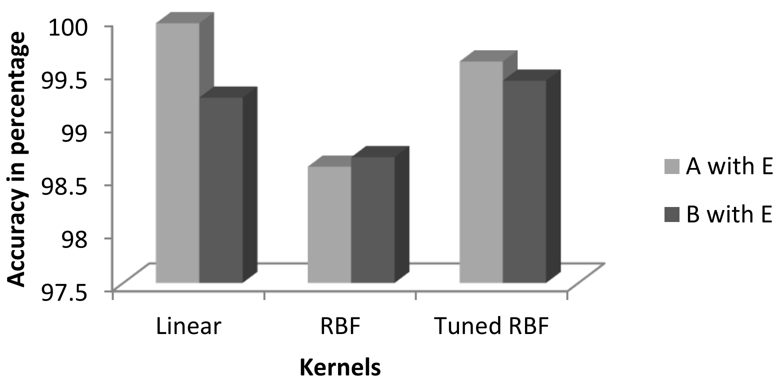


Figure 5: Binary classification results

The analysis has been carried out over the EEG signal and deals as it is. The frequency characteristics of the alpha, beta, theta and delta are invisible in this spatial domain. The exclusivity of this study describes that any segment of the signal from normal EEG and seizure EEG can be trained and tested efficiently. The signal features which are classified for seizure detection by using SVM classifier with the linear kernel have been obtained successful performance accuracy. Also the analysis proves that the RBF kernel gets good result by tuning the parameter sigma value.

CONCLUSION

The study works out well for any combination of the segment of normal and seizure signals for binary classification. Each segment of the EEG signal is compressed to 8 statistical and 4 GLCM features. These features are analyzed and classified by using two different SVM kernels. The study declares that the linear kernel is considered to be the best choice in this analysis. Further the parameter sigma in RBF kernel is tuned to the value 2 and compared with the default value 1. The analysis has proved that the RBF kernel achieves good results by tuning the sigma value. Considering the EEG signal directly for feature extraction is the limitation of this work. The frequency characteristics of EEG signals, namely alpha, beta, theta, delta and gamma are not analyzed. The signals are not decomposed based on their frequencies. The normal EEG includes alpha and beta, whereas the frequency of seizure is identified by slow waves (theta and delta). So it is necessary to analyze the EEG signals according to their frequency.The conventional learning systems like neural networks experience of their theoretical weakness, for example back-propagation algorithm generally converges only to local optimal solutions. At this point, SVM can afford a considerable enhancement. ANN is based on parametric models (finite dimensional model) which has the number of hidden layers with different sizes depend on number of features and it also has a bias parameter. But SVM is based on non-parametric model which is highly accurate and able to model complex non linear decision boundaries.

REFERENCES

[1] Adeli,H.,Zhou,Z., & Dadmehr,N. Journal of neuroscience Methods 2003; 123: 69-87.
[2] Andrzejak R.G., Lehnertz K., Mormann F., RiekeC.David P, & ElderC.E., Physical Review E2001; 64(6): 61907. http://epileptologie-bonn.de/cms/front_content.php?idcat=193
[3] Cortes,C., and Vapnik,V. Machine Learning 1995; 273-297.
[4] Fritz Albregtsen. Image Processing Laboratory. Department of Informatics, University of Oslo, 2008.
[5] Gandhi T, Panigrahi B.K & Anand S. Neurocomputing 2011; 74: 3051-3057.
[6] Gular I. and Ubeyli E.D. IEEE Transactions on information technology in Biomedicine 2007; 11: 117-126.
[7] Jurack V. Tsuzuki D and Dan.I. 10/20, 10/10 and 10/5 systems revisited: NeuroImage 2007; 34: 1600-1611.
[8] Long Han, Mark J, Embrechts, Boleslaw Szymanski, Karsten Sternickel, Alexander Ross. Computational Modeling and Simulation of Intellect 2011; 206-223.
[9] Misra U.K & Kalita J. published by Elsevier. A division of Reed Elsevier India Private Limited 2009.
[10] Nanthini,B.S. and Santhi,B. Journal of Applied Sciences 2014;.14: 1658-1661.
[11] Nanthini,B.S. and Santhi,B. Int. J.Signal and Imaging Systems Engineering 2015; 8: 2015, 28-38.
[12] Nicolaou N & Georgiou J. Expert Systems with Applications. 2012; 39: 202-209.
[13] Orhan U, Hekim M, Ozer M. Expert Systems with Applications 2011; 38: 13475-13481.
[14] Simon Haykin 1994. Prentice Hall PTR USA 1994. ISBN: 0023527617.
[15] Siuly, Yan Li, Peng-(Paul) Wen. Computer Methods and programs in Biomedicine 2011; 104: 358-372.
[16] Subasi A. Expert Systems with Applications 2005; 28: 701-711.
[17] Subasi A and Ercelebi E. Computer Methods and Programs in Biomedicine 2005; 78: 87-99.
[18] Subasi A. Expert Systems and Applications 2005; 31: 320-328.
[19] Ziqiang Li, Mingtian Zhou, Hao Lin, Haibo Pu. International Journal of Machine Learning & cybernetics 2014; 5: 425-434.