I J C T A, 9(17) 2016, pp. 8703-8708

Detection of Seizure in EEG Signal using Classical Pattern Recognition Tools

Pratibha Singh*

ABSTRACT

This paper is about the application of various pattern recognition techniques for seizure detection. The EEG signals of the subjects are obtained from publically available dataset. The features are then extracted which are based on Empirical mode decomposition(EMD). The training is done using a fixed percentage of training data or the cross validation. The result for cross validation is found better than the fixed percentage method.

Keywords: EMD, Features, k-Nearest Neighbor(k-NN), classifier

I. INTRODUCTION

The abnormal electrical activity causing mental disorder results in seizure. When it becomes recurrent processes then it causes phenomenon called epilepsy. The condition of seizure may vary from virtually un-noticeable state to the state of loss of consciousness and theconditions causing seizures are:Head injuries, Stroke, Electrolyte imbalance,Brain tumors, Very low blood sugar, Repetitive sounds or flashing lights, such as in video games, Medications, such as antipsychotics and some asthma drugs, Withdrawal from medications, such as Xanax, narcotics, or alcohol, use of drugs such as cocaine and heroin, Cancer andBrain infections, such as meningitis [1]. Electroencephalogram (EEG) continues plays a key role in diagnosis of seizure disorders [2] in patients because it is a convenient and relatively inexpensive way to demonstrate the physiological manifestations of abnormal cortical excitability that underlie epilepsy. However it suffers from the following drawbacks: While recording electrical activity by electrodes located on the scalp or surface of the brain, most of the times we get summation of excitatory and inhibitory postsynaptic potentials in apical dendrites of pyramidal neurons in the more superficial layers of the cortex.

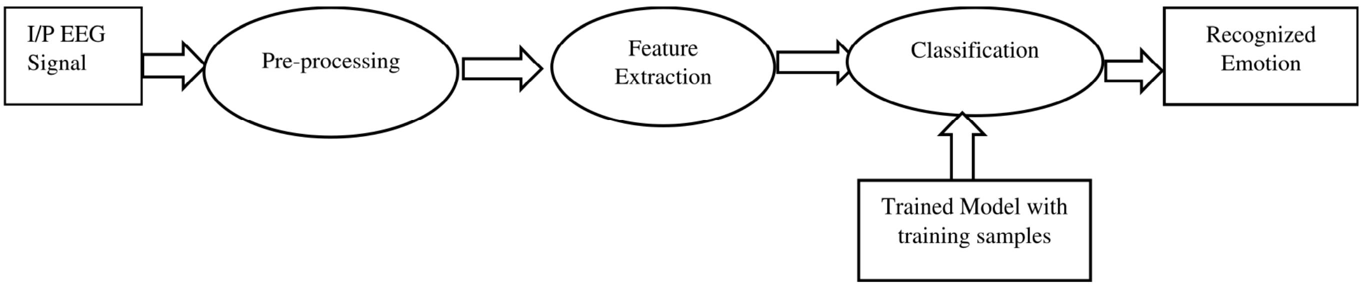
In Seizure some noticeable changes can be observed in EEG signal due to synchronous electrical activity of the neurons. The occurrence of spikes and sharp waves are some important characteristics present in EEG[3]. Srinivasan et al usedfeatures of spectral domainusing the Fourier transform forthe classification of epilepticseizure in EEG signals. They proposed an automated detection system for epileptic seizures using a version of Recurrent Neural network known as Elman network (EN)[4]. The methods based on Fourier transform analysis is utilizing the assumptions that the signal being experimented is stationary. But later on the EEGsignal was characterized to be belonging to a non-stationary process [5] by researchers. Sometime–frequencydomain based methods were developed for detection ofepileptic seizure from EEG signals. The methods based on time-frequency domain include the shorttime Fourier transform(STFT) [6], the wavelettransform[7], themulti-wavelet transform [8], the smoothed pseudo-Wigner–Ville distribution [9], and the multifractal analysis and wavelet transform[10]. The improved generalized fractal dimension hasbeen used for discriminating ictal EEG signals[11]. Recently,empirical mode decomposition (EMD) based methods for classification of ictal EEG signals have also been reported in literature [12] [13] [14] [15] [16]. Since the EMD based method is quite widely and successfully tested for seizure detection, this work is also using the EMD based features for seizure detection. In this work a model is proposed in which

* Assistant Professor, Electronics & Instrumentation Engg., IET DAVV, Indore, India, *E-mail: prat_ibh_a@yahoo.com*

various classification models are tested for the detection of seizure. The dataset of seizure had been obtained from BONN university [17]. The performance is measured in terms of percentage of error. The rest of the paper is organized as follows: section 1 gives the introduction, section 2 gives the proposed methodology, section 3 discusses about the classification methods, section 4 gives experimental results and conclusion is given section 5.

II. PROPOSED METHODOLOGY

The method for seizure detection is done using standard method of pattern recognition which is given in figure 1. Input for the system is the dataset of EEG signal. The dataset obtained from BONN University [17] has five classes defined as Z, N, O, S and F, out of which experiments were conducted for class F, N and S. Class F and N belongs to Non-seizure category while the Z class belongs to Seizure category.



The dataset is divided into training and testing set. Both the training and testing set undergo various preprocessing and feature extraction stages. For the extraction of features, EMD based second order difference plot area[13] was used. The training-testing pattern separation is done using cross-validation method. For detecting seizure the performance of various classifiers are tested namely Karhunen-Loève Mapping(klm), k-Nearest neighbor classifier, nearest mean classifier, Linear discriminant classifier, Quadratic discriminant classifier, Forward feature selection based on Nearest Neighbor, Forward Feature selection based on linear discriminant classifier.

2.1. Empirical Mode Decomposition[18]

The empirical mode decomposition is widely used for the signals which are nonlinear and nonstationary. It is an adaptive and data dependent method and does not require the condition like linearity and stationarity of the signal. This method is based on decomposition of non-linear and non-stationary signal x(t) into the sum of intrinsic mode function(IMF's). There are several feature extraction methods proposed using EMD[12][13][14][15], in all these methods firstly the EMD[18]is applied to obtain IMF's (Intrinsic Mode Function) of each signal. In this work features are calculated as the area of second order difference plot[13][14]on first 8 IMF functions. The method of obtaining IMF's from EMD is as follows:-

1) Detect the maxima and minima of the given data set
2) By connecting maxima and minima separately, generate upper and lower envelopes.
3) Calculate mean using equation (1)

$$\mathrm{a(t)} = \frac{[e_m(t)+e_l(t)]}{2} \tag{1}$$

where $e_m(t)$ is the maximum and $e_l(t)$ is the minimum value.

4) Extract IMF using equation (2)

$$h_1(t) = = \mathrm{x(t) - a(t)} \tag{2}$$

IMF'sare obtained by repeating above algorithm. Now the analytic signal z(t) of any real IMF y(t) is defined as

$$z(t) = y(t) + j\,HT(y(t))\,,\ (\text{ whereHT }(\mathrm{y(t)}) = \mathrm{y(t)} * \frac{1}{\pi t})$$

$$\mathrm{z(t)} = \mathrm{A(t)}e^{-j\emptyset(t)} \tag{3}$$

Where, A(t) = signal amplitude

Now EEG signal and their IMF functions are considered to be like a time series data and are represented in the form of phase space representation (PSR). The pattern of oscillatory signal is found to be elliptical in PSR. The area parameter as used by Pachori et al.[14]is used as features in this work.

III. CLASSIFICATION

The Baye's Theorem based classifier defines a rule by which instances can be sorted among various classes: an instance will be assigned to the class with the highest posterior probability given that it has the characteristics of the measurement vector **x**. This is the Baye's optimum, or maximum likelihood decision rule:

$$x \in \omega_i iff \frac{p(\omega i|x)p(\omega i)}{p(x)} \geq \frac{p(\omega i|x)p(\omega i)}{p(x)} forall\omega i \tag{4}$$

Classifier's accuracy:

$$Accuracy = P(\ correct\ ) = \sum_{i=1}^{N} \int_{Di} p(x|\omega i)p(\omega i)dx \tag{5}$$

$$Error = P(error) = \sum_{i=1}^{N} \int_{Di} \sum_{\substack{j=1 \\ j \neq i}}^{N} p(x|\omega j)p(\omega j)\, dx \tag{6}$$

The assumption of accurate assumption of conditional probability is valid for low dimensional problem, But error occurs when the data dimension increases. The most spread density function to estimate the conditional probability is the normal distribution given by

$$p(x|\omega i) = N(\mu_i, \textstyle\sum_i)$$

$$p(x|\omega i) = \frac{1}{(2\pi)^{\frac{n}{2}}\sqrt{|\sum_i|}} \exp\left[-\frac{1}{2}(\mathrm{x} - \mu_i)^{\mathrm{T}} \textstyle\sum_{\mathrm{i}}^{-1}(\mathrm{x} - \mu_i)\right] \tag{7}$$

Where $\mu_i \in \Re D$ is the mean of i^{th} class and

Σ_i is DxD covariance matrix, D is the dimension of feature vector. The estimation method for the class conditional density can be of two types parametric and non parametric. Parametric method is based on assumption of certain distribution function defined by set of parameters such as Gaussian distribution function. The parameters of the distribution are obtained by Maximum Likelihood decision rule given by equation (1). The LDC (Linear Bayes Normal Classifier), Quadratic Bayes Normal Classifier(QDC)methods discussed in following subsection are belonging to parametric methods. In nonparametric methods the pre-defined distribution function is not used but the training data itself used to estimate the distribution. K-Nearest Neighbor classifier belongs to this category.

3.1. Linear and Quadratic Classifiers

Density based classifiers can be either a Linear classifier or quadratic classifiers. They are designated so because of their type of discriminant functions. Thus any set of linear functions $g_i : \Re^n \rightarrow \Re$, i=1,...c. 'c' is the number of classes.

$$g_i(x) = w_{i0} + w_i^T x, \qquad x, w_i \in \Re^n \text{and} w_{i0} \in \Re \tag{8}$$

canwork as a linear classifier. These are the minimum-error (Bayes) classifiers defined for normally distributed classes having equal covariance matrices. LDC model can be simplybuild from give data and is reasonably good even when the classes do not have normal distributions. The discriminant functions can be obtained from the posterior probabilities $p(\omega i|\mathbf{x})$ by applying monotonic transformation.The resulted model will be optimal one in terms of error. The set of discriminant function in terms of probability can be defined using equation (9).

$$g_i(x) = \log\big(P(\omega i)p(x|\omega i)\big), \qquad i=1,.............c \tag{9}$$

The equation 9 can also be re-written as

$$g_i(x) = \log(P(\omega i)) + \log\left(\frac{1}{(2\pi)^{\frac{n}{2}}\sqrt{|\Sigma_i|}}\exp\left[-\frac{1}{2}(\mathrm{x}-\mu_i)^{\mathrm{T}}\Sigma_{\mathrm{i}}^{-1}(\mathrm{x}-\mu_i)\right]\right)$$

$$=\log(P(\omega i)) + \frac{n}{2}\log(2\pi) + \frac{1}{2}\log(|\Sigma_i|) + \frac{1}{2}(\mathrm{x}-\mu_i)^{\mathrm{T}}\Sigma_{\mathrm{i}}^{-1}(\mathrm{x}-\mu_i) \tag{10}$$

Assuming that all class-covariance matrices are the same in the above equation, that is,

$$\Sigma_i = \Sigma \text{ and } p(x|\omega i) \sim \mathrm{N}(\mu_i, \Sigma).$$

When the classes are normally distributed and covariance are class specific then the classifier is becomesQuadratic Discriminant Classifier(QDC). The Discriminant function is given by the eq. (11)

$$g_i(x) = w_{i0} + w_i^T x + x^T W_i x \tag{11}$$

where $w_{i0} = \log(P(\omega i)) + \frac{1}{2}\mu_i^T \Sigma_i^{-1}\mu_i + \frac{1}{2}\log(|\Sigma_i|)$, $w_i = \Sigma_i^{-1}\mu_i$ and $W_i = -\frac{1}{2}\Sigma_i^{-1}$

The estimates of the parameters for LDC and the quadratic discriminant classifier (QDC) are calculated from data. Let Ni be the number of objects in our data set Z from class ω_i, $i = 1, ...; c$, and $l(z_j) \in \Omega$ be the class label of $z_j \in Z$. The means are obtained by

$$\mu_i = \frac{1}{N_i}\sum_{l(z_j)=\omega i} z_i \tag{12}$$

and the covariance matrices, by

$$\Sigma_i = \frac{1}{N_i}\sum_{l(z_j)=\omega i}(z_j - \mu_i)\,(z_j - \mu_i)^T \tag{13}$$

The common covariance matrix for LDC is obtained as the weighted average of the separately estimated class-conditional covariance matrices calculated as $\Sigma = \frac{1}{N}\sum_{i=1}^{c} N_i\,\Sigma_i$.

3.2. Nearest Mean Classifier (NMC)

Nearest Mean Classifier or Nearest Centroid Classifier, in machine learning, is used as a classification model which assigns the observations, the label of class of the training samples whose centroid or mean is very closed to the observation.NMC is a plain nearest mean classifier for which the assigned classes are sensitive to feature scaling and insensitive to class priors.

3.3. Nearest neighbor Classifier(k-NN)

It is the well-known classifier where the class label to a test sample is assigned on the basis of its closest training sample. The k-NN rule is considered to be quite simple classification rule but is very computationally intensive.This method does not require any prior information. Let $T = \{s_1; s_2; : : : ; s_n\}$denote the set of n-labeled training samples. Each sample is a d-dimensional vector. Let $s_i$be the training sample nearest to a given test sample t in terms of some metric or distance function. The nearest neighbor rule for classifying t is to assign it to the class to which $s_i$belongs. The metric used for the present work is the Euclidean distance.

k-Nearest Neighbor

K- Nearest neighbor classifier is based on selecting a set of labeled prototype for each class. The classification is performed using a similarity measure between the unknown test sample and these set of labeled prototypes. Let $T = \{s_1; s_2; : : : ; s_n\}$denote the set of *n*-labeled training prototypes. Given a test sample *t*, let $R = \{r_1; r_2; : : : ; r_k\}$ be a set of the *knearest* training prototypes to *t* in terms of some metric. The *k*-nearest neighbor rule is to assign the sample *t* to the class whose frequency of occurrence is most among k-nearest training samples. Again the metric used is the Euclidean distance. The values of *k* used in this work is 1. The advantages of using nearest neighbor classifier are: it does not require learning, works well even for few training samples, good performance for lower dimensional space etc. The drawbacks of k-NN are slow performance and suffers from curse of dimensionality problem[19].

IV. EXPERIMENTAL RESULTS

In this paper experiments were performed on EEG dataset available online. After applying steps of the methodology explained in previous subsections, the features size we get was 8. Then from this dataset made up of observations stored in feature form,the data will be selected for training purpose and for testing purpose. There are two methods by which data is broken down in training and testing set: k-fold cross validation and by changing the percentage of training. Then the results were obtained using 5-fold cross validation and the error rate are tabulated in Table 2.

Table 1
Percentage error using k-fold cross validation

S. No.	*1st fold*	*2nd fold*	*3rd fold*	*4th fold*	*5th fold*	*Average Error*
Klm	25	35	28.3	28.3	33.33	29.98
Feature self(NN)	23.3	23.3	23.3	18.3	23.33	22.3
Feature self(ldc)	16.7	20	18.3	16.7	18.3	18
ldc	15	20	16.7	16.7	15	16.68
1-NN	1.7	5	0	3.33	0	2
qdc	3.3	0	3.33	3.33	0	1.99

Various classification tools for recognition are used from the PRTOOL pattern recognition toolbox[20]. The first column indicates the type of classifier used. Experimental result of cross validation is given in Table1 and the error rate for fixed training and testing size is given in table Table 2.

Table 2
Percentage error using fixed percentage of training

S. No.	*80 % training data*	*70% training data*	*60 % training data*	*50 % training data*
Klm	26.7	25.6	27.5	28.7
Feature self(NN)	18.3	20	20.8	24.7
Feature self(ldc)	15	18.9	15.8	18.7
ldc	15	17.8	17.5	16
1-NN	1.7	1.1	0.8	1.3
qdc	0	1.1	1.7	4

V. CONCLUSION

In this study the improved performance is observed for epilepsy seizure detection. While using fixed percentage of tainting data the result obtained for 70 % training data was best. For the experiments utilizing cross validation, the result of quadratic discriminate classifier was found to be the best. Best result as low as 0% error rate is observed for quadratic classifier using 80% training data.

REFERENCES

[1] D. team, "WebMD Seizure Disorders Directory," WebMD LLC, Office of Privacy, 1201 Peachtree Street NE, 400 Colony Square, Suite 2100, Atlanta GA 30361, 2006-2016. [Online]. Available: http://www.webmd.com/brain/seizure-disorders-directory.

[2] L. D. Iasemidis, D.-S. Shiau, W. Chaovalitwongse and J. C. Sackellares, "Adaptive epileptic seizure prediction system," *IEEE Transactions on Biomedical Engineering,* vol. 50, no. 5, pp. 616-627, 2003.

[3] S. Altunay, Z. Telatar and O. Erogul, "Epileptic EEG detection using the linear pre-diction error energy," *Expert Systems with Applications 3,* vol. 37, no. 8, pp. 5661-5665, 2010.

[4] V. Srinivasan, C. Eswaran and N. Sriraam, "Artificial neural network based epilep-tic detection using time-domain and frequency-domain features," *Journal of Medical Systems,* vol. 29, p. 647–660., 2005.

[5] B. Boashash, M. Mesbah and P. Colditz, "Time frequency detection of EEG abnor-malities, in Time–Frequency Signal Analysis and Processing: A ComprehensiveReference," Elsevier, Oxford , 2003,, pp. 663–670, ch.15, article 15.5..

[6] A. W. K. S. K. C. R. Schuyler, "Epileptic seizure detection," *IEEE Engineering in Medicine and Biology Magazine,* pp. 74-81, March 2007.

[7] H. Ocak, "Optimal classification of epileptic seizures in EEG using wavelet anal-ysis and genetic algorithm,," *Signal Processing,* vol. 7, p. 1858–1867, 2008.

[8] L. Guo, D. Rivero and A. Pazos, "Epileptic seizure detection using multiwavelet trans-form based approximate entropy and artificial neural networks," *Journal of Neuroscience Methods,* vol. 193, no. 1, pp. 156-163, 2010.

[9] A. Tzallas, M. G. Tsipouras and D. Fotiadis, "Epileptic seizure detection in EEGsusing time–frequency analysis," *IEEE Transactions on Information Technologyin Biomedicine,* vol. 13, no. 5, pp. 703-710, 2009.

[10] R. Uthayakumar and D. Easwaramoorthy, "Epileptic seizure detection in EEG signalsusing multifractal analysis and wavelet transform," *Fractals* , vol. 21, no. 2, June 2013.

[11] D. Easwaramoorthy and R. Uthayakumar, "Improved generalized fractal dimensionsin the discrimination between healthy and epileptic EEG signals," *Journal ofComputational Science,* vol. 1, pp. 31-38, March 2011.

[12] V. Bajaj and R. Pachori, "Classification of seizure and nonseizureEEG signals using empirical mode decomposition," *, IEEETransactions on Information Technology in Biomedicine,* vol. 16, no. 6, p. 1135–1142, 2012.

[13] V. Bajaj and R. Pachori, "Epileptic seizure detection based onthe instantaneous area of analytic intrinsic mode functionsof EEG signals," *Biomedical Engineering Letters,* vol. 3, no. 1, pp. 17-21, 2013.

[14] R. Pachori and S. Patidar, "Epileptic seizure classification in EEG signals using second-order difference plot of intrinsic mode functions," *computer methods and programs in biomedicine,* vol. 1, no. 3, p. 494–502, 2014.

[15] R. Pachori and V. Bajaj, "Analysis of normal and epilepticseizure EEG signals using empirical mode decomposition," *Computer Methods and Programs in Biomedicine,* vol. 104, no. 3, p. 373–381, 2011.

[16] R. P. V. Bajaj, "EEG signal classification usingempirical mode decomposition and support vector machine," in *International Conference on Soft Computingfor Problem Solving, AISC* , Roorkee, India, 2011.

[17] R. Andrzejak, K. Lehnertz, C. Rieke, F. Mormann, P. David and C. Elger, "epileptologie," 2001. [Online]. Available: http://epileptologie-bonn.de/cms/front_content.php?idcat=193&lang=3. [Accessed august 2015].

[18] N. E. Huang, Z. Shen and S. R. Long, "The Empirical Mode Decomposition and the Hilbert Spectrum for Nonlinear and Non- Stationary Time Series Analysis," in *Proc. R. Soc.*, London, 1998.

[19] W. Powell, Approximate Dynamic Programming:Solving the Curses of Dimensionality, 1st edition ed., Wiley-Interscience, 2007.

[20] R. Duin, "PRTools 3.0, A Matlab Toolbox for Pattern Recognition, Delft University of Technology, .," 2000.