ISSN: 2321-2152

IJMECE

International Journal of modern electronics and communication engineering

E-Mail
editor.ijmece@gmail.com
editor@ijmece.com

www.ijmece.com

ISSN2321-2152www.ijmece .com
Vol 4, Issue.1Jan 2016

Statistical Wavelet Features, PCA, and SVM Based Approach for EEG Signals Classification

Mr. Rakeshkumar

***Abstract*—** Electroencephalography is the study of the electrical signals generated by neural activity in the human brain. In this study, we offer a method for efficiently classifying EEG signals that is both automated and robust. The suggested method is applied to the EEG signal in order to categorize it as either indicative of an epileptic seizure or not. We begin the suggested method by using Discrete Wavelet Transform (DWT) to extract features by breaking down the EEG data into frequency bands. These characteristics are input to Principal Component Analysis (PCA) and are derived from the information included in the details and approximation coefficients of DWT sub-bands. PCA is used to reduce the number of features used in the classification, while SVM is used to derive support vectors for use in the classification. The tests are conducted using a representative sample of the population. The end outcome of the categorization process achieves a high degree of accuracy

***Keywords*—** Electroencephalogram, Principal Component Analysis, Support Vector Machine, Discrete Wavelet Transform, and Discrete Wavelet Transform.

I. Introduction

~~The~~ field of ELECTROENCEPHELOGRAPHY investigates brain activity through the use of electrical recordings. Brain activity results in the generation of electrical impulses as early as the 17th week of gestation. The brain's electrical impulses are a reflection of mental processes and physiological states. Research fields as diverse as seizure detection and prediction, brain-computer interfacing, the study of psychiatric disorders and dementia, and sleep signal analysis can all benefit from a thorough examination of EEG signals. The EEG signal waves are further subdivided into five distinct sub-bands depending on frequency ranges for better understanding of human behavior. Delta ()(Range 0.5-4Hz), Theta ()(Range 4-8Hz), Alpha ()(Range 8-13 Hz), Beta ()(Range 13-30 Hz), and Gamma ()(Range 30-45 Hz) are the names given to these frequency bands in order from low to high.[1].

Seizures may be visually distinguished from other sources of noise in an EEG test by their distinctive, spiky waveforms. While most artifacts are non-stationary and irregular in shape, seizure signals have a periodic, conspicuous spiky pattern. However, when the uniqueness of the recorded EEG pattern is taken into

R. K. Chaurasiya may be reached at +91-91 65 971639 or by email at r.k.chaurasiya@nitrr.ac.in at the National Institute of Technology in Raipur, India. He works in the department of electronics and telecommunications there.
N. D. Londhe may be contacted at nlndhe.ele@nitrr.ac.in, which is the address of the Department of Electrical Engineering at

Assistant Professor,Department of Law , Ch.DeviLalUniversity,Sirsa(HR) Adv.rakesh48@gmail.com

the National Institute of Technology in Raipur, India.

When we last saw S. Ghosh, he was working at the NIT in Rourkela. Currently he may be found working in the Electrical Engineering Department at NIT Raipur in India. (Sghosh.ele@nitrr.ac.in is his email address).

It is quite challenging for a person to view and understand the real functioning of the brain by visual examination alone due to the mapping of data recorded by insertion of electrode into different parts of the scalp. As a result, there is a growing need for a cheap, user-friendly, and completely automated EEG-based seizure detection system.

In this study, we offer a method for identifying epileptic seizures based on statistical features. EEG signals are Discrete Wavelet Transformed (DWT) to extract statistical characteristics. Additionally, epilepsy and normality are classified using Support Vector Machine (SVM). Principal Component Analysis (PCA) was performed on the normalized feature matrix to decrease the time and space complexity and to eliminate redundancy in the observed characteristics.

II. RELATED WORK

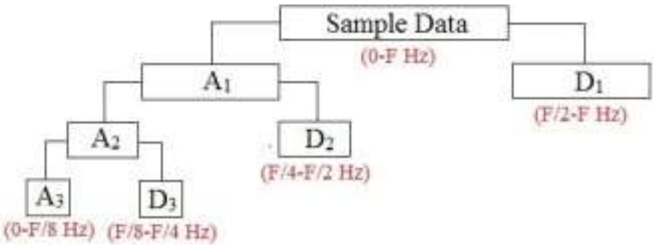
In 1875, English scientist Richard Caton discovered the electrical impulses for brain function. EEG research on the human brain was initiated by Hans Berger in 1920 [2]. The Greek root of the term epilepsy means "to seize" or "to attack." Early descriptions of epilepsy may be traced back to ancient Indian medicine (4500-1500 BC), when it was known as apasmara (literally "loss of consciousness"). A Babylonian tablet housed in London's British Museum also provides specifics about epilepsy and its treatment [1]. According to Kaufman, the seizures were caused by irregular current flow [3].

Visual evaluation of EEG data by highly trained electroencephalographers was the foundation of most 20th-century epilepsy analysis procedures. However, in the last two decades, thanks to advancements in signal processing and pattern recognition, a variety of automated algorithms for detecting epileptic seizures have been created [6, 9].

The frequency domain information is supplied at the expense of time domain information, such as the amplitude distribution and the EEG pattern, making spectral analysis based feature extraction a poor choice for EEG classification. This has led to the usage of Discrete Wavelet Transform (DWT) and other time- and frequency-domain based feature extraction techniques in recent studies [4–6]. The ability to apply DWT to the study of non-stationary signals, such as EEG [7], [8], is another area in which it excels over spectral analysis. Recent work using a Hilbert-Huang Transformed (HHT) technique has been reported by Kai Fu et al. [9].

III. DATA-SET FOR EXPERIMENTAL ANALYSIS

Over the past few years, the vast majority of epilepsy seizure detection research has relied on the publicly available data outlined in [10]. In order to compare our findings to those of other studies, we are utilizing the same benchmark database. Different participants' data are used to compile the database, which is then organized into five sets (A-E) of 100 single-channel EEG samples each. At the time of data collection, the participants in each



dataset (A-E) were in the following states of mind:

Five healthy volunteers, A with eyes open and B with eyes closed, showing the calm awake awake condition.

C, D: Measuring what the patients were doing during seizure-free periods of EEGs from five people who had successfully managed their seizures and had accurate diagnoses.

Set E: Includes seizure activity (from the same individuals as B and C).

The data set was obtained with a standardized 10-20 electrode placement technique and a 128-channel amplifier setup. The information was then sampled and digitally converted at a rate of 173.61 samples per second with a resolution of 12 bits. Band-pass filter with 0.50-40 Hz (12 dB/oct) applied since the relevant data is only present in the,,, and sub-bands. Classification in this study was performed using two datasets, A and E. While dataset E comprises samples from people with verified epilepsy (Class I), dataset A contains samples from people without epilepsy (Class II).

IV. PROPOSED ALGORITHM

V. 10000595Table I lists each individual step of the proposed method. Following the steps in Table I, we extract the DWT coefficients from a random EEG sample. The DWT coefficients are parsed for features, which are then inserted into the appropriate column of a feature matrix. Keep in mind that if you have 100 EEG samples, your feature matrix will contain 100 columns and however many features you extracted for rows. All of the EEG samples go through the same process. The resultant feature matrix is normalized and sent on to principal component analysis (PCA) for dimension reduction. The dimension-reduced feature matrix is then classified using binary SVM.

VI.

Fig. 1 3-Level wavelet decomposition of the sample data signal having 0-F Hz frequency range. The signal is decomposed into detail coefficients D_1-D_3 and approximation A_3. The frequency range covered in different decompositions and approximation is shown in the bracket

TABLE I
STEPWISE DETAILS OF THE PROPOSED ALGORITHM

1. i = 1
2. **for** i< = size of the data set, **do**
3. Decomposition the ith EEG sample using5-level DWT.
4. Extract the statistical wavelet features from DWT coefficients, and put in ith column of feature matrix *Ftr_Mat*.
5. **end for.**
6. Normalization the *Ftr_Mat* (feature wise).
7. PCA on *Ftr_Mat* for dimension reduction.
8. Train the SVM and derive the support vectors.
9. Apply SVM on test data for Classification.
10. Measure the accuracy obtained by SVM classification.

A. *Feature Extraction Using DWT*

(1) For study of constant signals, the Fourier transform and related spectrum analysis methods are widely employed. It is not

advised to apply the Fourier transform directly to non-stationary signals like EEG. Therefore, the suggested approach makes use of a time-frequency analysis method based on a wavelet transform.

(2) Multi-level wavelet decomposition of EEG recordings yields information at fine-grained frequency-specific resolutions [11]. Three-level wave decomposition is shown in Fig. 1.

(3) The criteria for choosing the decomposition level and fundamental wavelet type are situationally dependent. The frequency range of 0-50 Hz is of importance for feature extraction from EEG recordings. Therefore, level 5 is selected as the breakdown granularity. The performance of the SVM classification was evaluated after experimenting with several wavelet types. The Daubechies wavelet was selected as the filter wavelet since it was shown to have the best fit for the EEG data. Sample A EEG signal, D1–D5, and A5 approximation are shown, decomposed, and discussed in Fig. 2. For feature extraction, we choose for the D3-D5 decomposition and the A5 approximation based on the relative frequency of the features of interest.

(4)

(5) The time and frequency representation of the EEG signals is obtained by extracting the wavelet coefficients. The following statistical characteristics are derived from these coefficients:

(6) Feature 1 to 4 consists of the mean of the absolute values of the approximation (A_5) and details (D_3-D_5).

[Ftr(1), Ftr (2), Ftr (3), Ftr (4)] = [mean(abs(A_5)),mean(abs(D_5)),mean(abs (D_4)),mean(abs(D_3))],

Here Ftr is the feature vector for one EEG sample.

(7) Feature 5 to 8 consists of the average of the square of the second order norm (equivalent to average power of discrete signals) of the approximation and details.

(8) Feature 9 to 12 consists of the median of the actual values of the approximation and details.

(9) Feature 13 to 16 consists of the standard deviation of the coefficients of the approximation and details.

(10) Feature 17 to 20 consists of the kurtosis; feature 21 to 24 consists of the skewness; and feature 25 to 28 consists of the entr

(11) opy of the coefficients of the approximation and details.

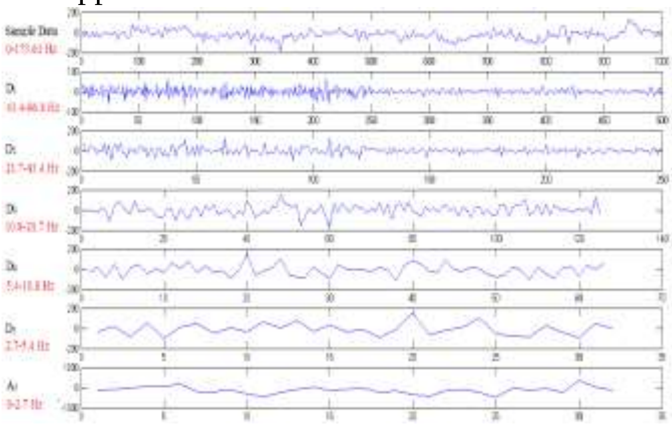


Fig. 2 5-Level wavelet decomposition of sample data (from set A) signal of 0-173.61 Hz. D_1-D_5 are details and A_5 is approximation. (For clear visibility only 1000 initial samples taken from 4097 samples of the sample data and axis are *not* equalon the sub-plots)

(12) Feature 29 to 31 consists of the ratio of the absolute mean

values of

adjacent sub-bands i.e.

approximation and

detail. The first three

feature sets

(1-3) represent the

frequency distribution of the signal and the other three (4-

to perform more accurately on unknown data [13], [14].

perform well

when it is applied for the data outside the training set. The notion of maximizing the *margin* between the

support vectors is at the heart of the SVM classifier, in order

classifier, in order

6) represents the variation with respect to frequency distribution. It is clear that feature vector for one EEG sample consists of 31 features. As each data set (A to E) consists of 100 EEG samples, the feature matrix for each

Consider the hyper-plane in (1):

$$w^T x + w_O = 0 \quad (1)$$

data set is of the dimension 31×100.

B. Principal Component Analysis (PCA):

PCA is well-established and the most widely used method

The *margin* is the Euclidian distance $1/\|w\|$ between the two parallel hyper-planes (support vectors) described in (2):

for dimension reduction. PCA allows to represent a *d*-

$$w^T x + w_O = 1, \text{ and } w^T x + w_O = -1 \quad (2)$$

dimensional data into a lower dimensional space (say *l*, where *l*<*d*). The PCA reduced data set is the *best* representation of the *d*-dimensional data into *l*-dimensional space (best in terms of minimum squared-error-distance).

In the process of minimizing the squared-error distance between the actual data and reduced data, one can derive the

Let x_i are training points, with respective classes $y_i \in \{-1,1\}$, i=1,2,….,N for a 2-class classification problem. The task is to optimize for minimum training error and maximum separating margin between hyper-planes of (2). SVM classifier solves this task by solving the optimization problem of (3):

method to reduce the *d*-dimension data into *l*-dimensional data

$$\text{Minimize } L(w, w_O, £) = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N} £_i \quad (3)$$

through PCA. First the *d*-dimensional mean μ and *d*×*d* dimensional covariance matrix S are computed for original *d*-dimensional data set. Next, *d* eigen values are calculated and

Subjected to

$$w^T x + w_O \geq 1 - £_i, \quad \text{if } y_i \in 1$$

are sorted in decreasing order. Say these eigen values (in

$$w^T x + w_O \leq -1 + £_i, \text{ if } y_i \in -1$$

decreasing order) are $\lambda_1, \lambda_2 \ldots. \lambda_d$ and the corresponding eigen vectors are $e_1, e_2 \ldots\ldots e_d$. (all the eigen vectors $e_1, e_2 \ldots\ldots e_d$ are mutually orthogonal). Subsequently, the first *l* eigen vectors

and

$$£_i \geq 0 \quad (4)$$

$e_1, e_2 \ldots\ldots e_l$, which correspond to largest *l* eigen values λ_1,

For the present work, involving two class classification

$\lambda_2 \ldots. \lambda_l$ are chosen as natural basis for projecting the *d*-

(Epileptic seizure or not), we first learn the classifier equation

dimensional data in *l*-dimensional space. A good value of *l* is decided by the fact that there is a significant comparative difference between l^{th} and $(l+1)^{th}$ eigen value. The more details and mathematical analysis of PCA can be found in [12] [13].

C. Support Vector Machine

The idea of SVM is originated from the idea of controlling the *generalizing capabilities* of machines for automation. The performance of a classifier must be generalized, i.e. it should

(similar to (1)) by solving the optimization problem of (3) with constrains of (4), using half of the feature vectors from *Ftr_Mat* as training data. Then the separating hyper-plane is used to classify the remaining feature vectors of *Ftr_Mat*. The pictorial representation of learning the SVM classifier from the training data of different classes is shown in Fig. 3.

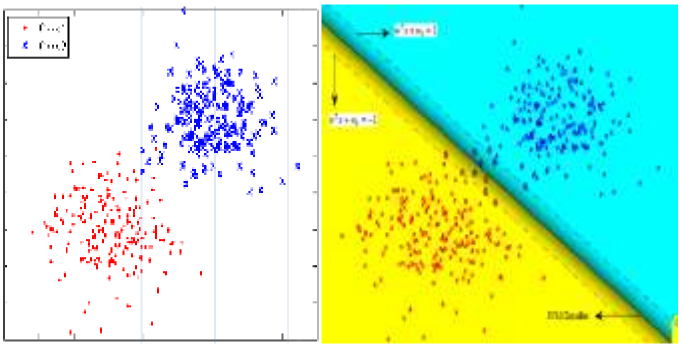


Fig. 3 (a) Distribution of 2-dimensional data set of two different classes (b) Support vectors (dotted lines) and SVM classifier (solid line) learnt to optimize for minimum training error and maximum separating margin between hyper-planes

VII. EXPERIMENTAL RESULTS

In the presented work, we have used pattern recognition approach for EEG signal classification. EEG samples were

The accuracy is calculated using:

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+TN+FN} \times 100\% \quad (8)$$

decomposed into sub-band using 5-level DWT with

Daubechies wavelet. Thirty-one different Statistical features were extracted from the details and approximation sub-bands. This feature extraction process was repeated for all the sample signals of set A and set E. A 31×200 sized *Ftr_Mat* was formed after adding all the feature vectors to it. Here each row

Table II summarizes the values of accuracy, sensitivity and specificity obtained after classification.

VIII. CONCLUSION

In this study, we offer a pattern recognition method where each column is assigned to a specific characteristic that may be used to identify an epileptic episode. The basis of the suggested method is

correlates to a certain EEG recording. After decomposing DWT into sub-bands, statistical characteristics such as mean, power, standard deviation, kurtosis, skewness, entropy, and median were extracted and used to normalize the rows of Ftr_Mat between 0 and 1. The collected wavelet features are normalized, and then sent to principal component analysis (PCA) for where xi = feature value in the ith row, xmin = dimension reduction, to get rid of any redundant features. The feature dimension is reduced by taking the minimum value in that row and replacing it with xmax= the highest value in that row.

To further minimize the time and space complexity, PCA helped reduce the dimensions of the collected features to 7. To divide the records into two groups—those with and those without epileptic seizures—PCA is used in conjunction with SVM classification. We were left with a 7200 after dimension reduction.

With Ftr_Mat, the detection ratio and precision both reach unprecedented heights. After implementing the recommended classification strategy on real data, we used 100 of the 200 feature points (50 from each class) to train the SVM classifier. The reliability of the SVM classifier was evaluated using the remaining feature points.

Captured brainwave activity (EEG) data. The system's high degree of accuracy makes it an ideal aid for automatic categorization.

The classification accuracy of EEG data may be evaluated in this manner.

The future of the matrix may lie in the development of a specialized hardware configuration and a user-friendly interface, both of which increased sensitivity (True Positive Ratio- TPR) and specificity (True Negative Ratio- TNR). The formula employed in the proposed study is described by Equations (6) and (7).

determining sensitivity and specificity by comparing the number of True Positive (TP) and True Negative (TN) results with the number of False Positive (FP) and False Negative (FN) results.

[1]

REFERENCES

Saeid Sanei and J.A. Chambers. *EEG Signal Processing*. John Wiley & Sons, 2007.

$$\text{Sensetivity} = TPR = \frac{TP}{TP+FN} \times 100\%$$

A. Massimo, "In Memoriam Pierre Gloor (1923–2003): an appreciation". Epilepsia, vol.-45(7), July 2004, page-882.

M. A. B. Brazier. "A History of the Electrical Activity of the Brain".

$$\text{Specificity} = TNR = \frac{TN}{TN+FP} \times 100\%$$

TABLE II

The First Half-Century, Macmillan, New York, 1961.

M. D. Alessandro, R. Esteller, G. Vachtsevanos, A. Hinson, A. Echauz, and B.Litt. "Epileptic seizure prediction using hybrid feature selection

THE VALUES OBTAINED FOR PERFORMANCE MEASURING

PARAMETERS

Parameter Numerical Value Obtained

over multiple intracranial EEG electrode contacts: A report of four patients". IEEE Transactions on Biomedical Engineering-2003. vol.-50 (5), pp.-603–615.

Sensitivity	Specificity	Accuracy
100.0%	99.50%	99.75%

[5]B. P. Marchant. "Time–frequency analysis for biosystem engineering".
2003, volume 85, issue 3, pages 261-281, Biosystems Engineering.Author: A. Subasi; Title: "EEG Signal Classification Using Wavelet Feature Extraction and a Mixture of Expert Model" Pages 1084–1093 in the 2007 volume of Expert Systems with Applications.

"EEG signal classification using PCA, ICA, LDA, and support vector machines" [6], written by A. Subasi and M. Ismail Gursoy. Pages 8659–8666 in the 2010 volume of Expert Systems with Applications.

"Epileptic seizure detection using dynamic wavelet network," by A. Subasi, appeared in volume 29 of Expert Systems with Applications in 2005, pages 343–355.

In "Classification of seizure based on the time-frequency image of EEG signals using HHT and SVM," K. Fu, J. Qu, Y. Chai, and Y. Dong discuss how to do just that. 2014, volume 13, pages 15–22, Biomedical Signal Processing and Control.

"Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: dependence on recording region and brain state" [9], R.G. Andrzejak, K. Lehnertz, F. Mormann, et al. Review of Physics Education 64(6):061907, 2001.

H. Adeli, Z. Zhou, and N. Dadmehr [10]. "Wavelet transform analysis of electroencephalogram data from a patient with epilepsy." 69–87 in vol. 123 no. 1 of the Journal of Neuroscience Methods in 2003.

A study by R. O. Duda, P. E. Hart, and D.G. Stork [11]. John Wiley & Sons, 2001, Pattern Classification, Second Edition.

Theodoridis, S., and Koutroumbas, K. [12]. Elsevier/Academic Press, 2009. Pattern Recognition, Fourth Edition.

Reference: Sastry, P. S. "An Introduction to Support Vector Machines" [13]. This is a chapter from "Computing and Information Sciences: Recent Trends," edited by J.C. Misra. New Delhi: Narosa Publishing House, 2003.

Rahul K. Chaurasiya, who grew up in a remote part of Madhya-Pradesh, India, attended the Maulana Azad National Institute of Technology in Bhopal for his undergraduate education. After that, he became a graduate student at the Indian Institute of Science in Bangalore. In 2011, he went there and got his Master of Engineering in System Science and Automation.

He joined Brocade Communications in Bangalore as a senior software developer after finishing his Master of Engineering program there. He accepted a position as an assistant professor at the National Institute of Technology in Raipur, India. In addition, he is currently pursuing a doctorate from the same university. His current research interests include BCI, signal processing, and pattern recognition.