The Use of Statistical Wavelet Features, PCA, and Support Vector Machines for EEG Signal Classification

K.BALA, B.SAI NATH
Asso.Prof[1], Asst.Prof[2]
Department of Ece
SRINIVASA INSTITUTE OF TECHNOLOGY AND SCIENCE , KADAPA
Chennai- Hyderabad Bypass Road, Ukkayapalli ,kadapa -516002

Abstract**—**The study of the electrical signals produced by neural activities of human brain is called Electroencephalography. In this paper, we propose an automatic and efficient EEG signal classification approach. The proposed approach is used to classify the EEG signal into two classes: epileptic seizure or not. In the proposed approach, we start with extracting the features by applying Discrete Wavelet Transform (DWT) in order to decompose the EEG signals into sub-bands. These features, extracted from details and approximation coefficients of DWT sub-bands, are used as input to Principal Component Analysis (PCA). The classification is based on reducing the feature dimension using PCA and deriving the support- vectors using Support Vector Machine (SVM). The experimental are performed on real and standard dataset. A very high level of classification accuracy is obtained in the result of classification.

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

INTRODUCTION

LECTROENCEPHELOGRAPHY is the study of the electrical signals produced by brain. Production of electrical signals as
a result of neural activity of the brain starts as early as from the 17th week of prenatal development. Electrical signals generated by the human brain represent the thinking of the mind and the status of the body. The close study of these Electro-EncephaloGram (EEG) signals is useful in many research areas such as detection and classification of event related potentials, seizure detection and prediction, brain–computer interfacing, Study of mental disorders like psychiatric disorders and dementia, and sleep signal analysis. For better understanding of human behavior, the EEG signal waves are further divided in five major sub-bands based on the frequency ranges. These bands from low to high frequencies respectively are called delta (δ)(Range 0.5-4Hz), theta(θ)(Range 4-8 Hz), alpha (α) (Range 8-13 Hz), beta (β)(Range
13-30 Hz), and gamma (γ)(Range 30-45 HZ)[1].

The visual distinction of seizure from common artifacts within an EEG measurement is based on the shape and spikiness of the waveforms. A signal with seizure have a rhythmical and prominent spiky, whereas the most of other artifacts are non-stationary and randomly shaped. But considering the fact that the recorded EEG pattern is a special

mapping of signals captured by placement of electrode onto different regions of the scalp, it is extremely difficult for human being to observe and understand the actual behavior of the brain by merely visual inspection. Hence there is an ever increasing demand of easily accessible and fully automatic epileptic seizure detection system using EEG signals.

In this paper we propose a statistical feature based epileptic seizure detection system. Statistical features are extracted from Discrete Wavelet transforms (DWT) of EEG signals. Further, Support Vector Machine (SVM) is used for classification into two classes i.e. is epileptic and normal. In order to reduce the time and space complexity and to avoid redundancy in the observed features, we have applied Principal Component Analysis (PCA) on the normalized feature matrix.

I. RELATED WORK

The electrical signals for brain activity were first recorded by the English scientist Richard Caton in 1875. Hans Berger started the study of EEGs from human brain in 1920 [2]. *Epilepsy* is a Greek word, which means 'to seize or attack'. The very basic concepts of epilepsy can be found in ancient Indian medicine (4500–1500BC) as *apasmara*, which means "loss of consciousness". Babylonian tablet in the British Museum in London also gives the detailed knowledge about the epileptic disease and its cure [1]. Kaufman associated the epileptic attacks with abnormal electrical discharges [3].

Most of the epilepsy analysis methods developed in the 20th century were based on the concept of visual inspection of EEG signals by highly skilled electroencephalographers. However, with the advancement in the field of signal processing and pattern recognition, different automatic techniques of epileptic seizure detection have been developed in last two decades [6], [9].

Spectral analysis based feature extraction method provides poor results for EEG classification as the frequency domain information is provided at the cost of time domain information such as the amplitude distribution and EEG pattern. Hence, both time and frequency domain based feature extraction algorithms such as Discrete Wavelet Transform (DWT) are being used in current research [4]-[6]. The other advantage of DWT over spectral analysis is its suitability for analysis of non-stationary signals like EEG [7], [8]. Kai Fu et al. have recently published their work with Hilbert-Huang Transformed (HHT) based approach [9].

II. DATA-SET FOR EXPERIMENTAL ANALYSIS

In last few years, most of the researchers have used publiclyavailable data described in [10] for their research work in thefield of epileptic seizure detection. We are also using the same *bench-mark* database in order to compare our results with the results of previous research works. The database is prepared by taking inputs from different subjects and is divided into five sets (A-E) each containing 100 EEG samples recorded through single channel. The mental status of the subjects in each data set (A-E) at the time of data recording was as follows:

A, B: Five healthy volunteers, relaxed in an awake state with eyes open (A) and closed (B).

C, D: Activities measured during seizure free intervals of EEGs from five patients, all of whom had achieved complete seizure control and were correctly diagnosed.

E: Contains seizure activity (recorded from the same patients as for set B and C).

The data set were recorded using a 128-channel amplifier system and standardized 10-20 electrode placement scheme. After recording, the data were sampled and digitized at 173.61 samples per second using 12 bit resolution. As the useful information from the data can be found only in δ, θ, α, β, and γsub-bands, a band-pass filter with 0.50–40 Hz (12 dB/oct) wasapplied. In this study, we used the dataset A and E for classification, as only set E contains the samples from confirmed epilepsy (Class I), the data set A consists of sample from persons having no epilepsy (Class II).

III. PROPOSED ALGORITHM

The stepwise details of the proposed algorithm are given inTable I. As described in Table I, we pick one EEG sample at a time and find its DWT coefficients. The features from the DWT coefficients are extracted and appended in the corresponding column of a feature matrix. Note that for a set of 100 EEG samples, the feature matrix will have number of column = 100, and number of rows = the total number of extracted feature. The same procedure is repeated for all the EEG samples. The final feature matrix is normalized and passed for dimension reduction using PCA. Binary SVM classification is performed on the dimension reduced feature matrix for classification.

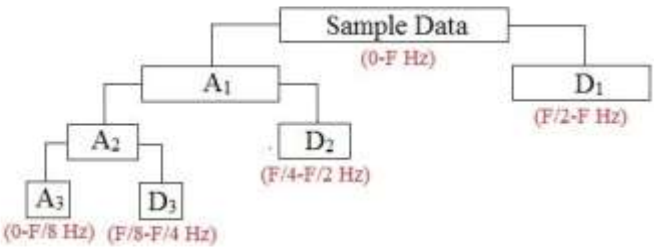


Fig. 1 3-Level wavelet decomposition of the sample data signal having 0-F Hz frequency range. The signal is decomposed into detail coefficients D_1-D_3 and approximation A_3. The frequency range covered in different decompositions and approximation is shown inthe bracket

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 putin 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

Fourier transform and other spectral analysis techniques are the popular tools used for analyzing stationary signals. However, for non-stationary signals like EEG, direct application of Fourier transform is not recommended. Hence, time-frequency analysis using wavelet transform have been adapted in the proposed work.

A multi-level wavelet decomposition of the EEG samples provides the information at different resolutions of the samples at different frequency bands [11]. Fig. 1 shows 3- level wave decomposition.

The selection of the level of decomposition and the type ofthe basic wavelet is a problem specific criterion. For extracting the features from EEG samples, the frequency range of interest is 0-50 Hz. So the level of decomposition is chosen to be 5. Different kinds of wavelets were tried and the accuracy of the SVM classification was measured. It was observed that the Daubechies wavelet suits the EEG signals more and hence it was chosen as filter wavelet. Fig. 2 shows an EEG sample signal from set A, its decomposition D_1-D_5 and approximation A_5. Considering the frequency of our interest, decomposition D_3-D_5 and approximation A_5 are chosen for feature extraction.

Extracted wavelet coefficients provides both time and
frequency representation of the EEG samples. Various statistical features are extracted from these coefficients as mentioned below:

(1) Feature 1 to 4 consists of the mean of the absolute valuesof 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.

(2) 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.
(3) Feature 9 to 12 consists of the median of the actual valuesof the approximation and details.
(4) Feature 13 to 16 consists of the standard deviation of thecoefficients of the approximation and details.
(5) Feature 17 to 20 consists of the kurtosis; feature 21 to 24consists of the skewness; and feature 25 to 28 consists of the entropy of the coefficients of the approximation and details.

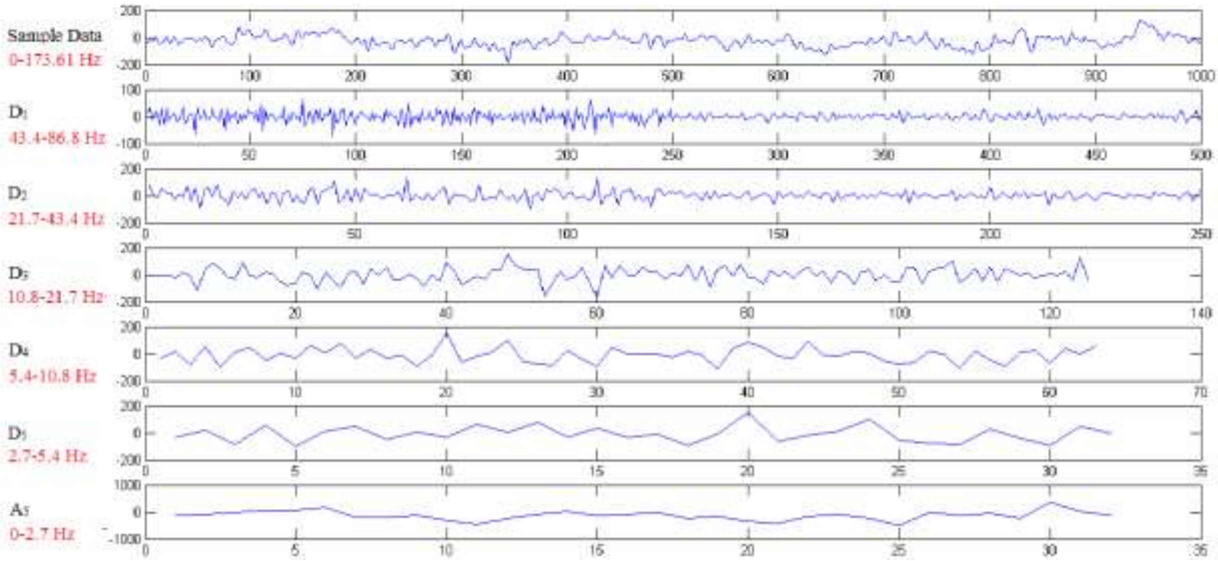


Fig. 2 5-Level wavelet decomposition of sample data (from set A) signal of 0-173.61 Hz. D_1-D_5 are details and A_5 is approximation. (For clear visibility only 1000 initial samples taken from 4097 samples of the sample data and axis are *not* equalon the sub-plots)

Sixth Feature: the absolute mean ratio, found in features 29–31.

Carry yourself admirably
application when used on information outside the relative importance to neighboring sub-bands, or the like. inference and approximation group of exercises Increasing the gap between the detail. Just the first three Indicators of Characteristics
The numbers (1-3) signify The SVM classifier relies on support vector machines. signal's frequency spectrum, and the other three (4-to get better results when working with ambiguous information [13], [14].

The frequency-dependent variation is denoted by (6. It can be seen that a single EEG sample has 31 characteristics that make up its feature vector. Each dataset (A–E) has 100 EEG samples, hence the feature matrix for each dataset

Take into account the hyper-plane in (1):

wTx + wO = 0

The data collection has 31 columns and 100 rows.

Secondary Factor Analysis (PCA)

It is commonly accepted that principal component analysis (PCA) is the most popular and effective technique.

The gap is the Euclidean distance between the two parallel hyper-planes (support vectors) defined by (2), which is equal to 1/||w||. on account of Scale
reduction.
Data collection using PCA enablessymbolizing a d- In both cases, 1 is the result of adding the wTx and wO values, while -1 is the result when switching signs. data of higher dimension (d) into one of lower dimension (let's assume l, where ld). In order to translate d-dimensional data into l-dimensional space, the principal component analysis reduced data set is optimal (best in terms of minimum squared-error-distance).

One may calculate the from the minimum squared-error gap between the original and filtered data.
For a 2-class classification issue, let xi be training points and yi be their respective classes (i=1,2,....,N). The goal is to find the optimal combination of training error and the margin of separation between the hyper-planes of (2). To do this, the SVM classifier optimizes the following three conditions:
technique for flattening d-dimensional datasets into l-dimensional ones

Lower the expression L(w, w, £) = 1 ||w |2 + C N £
using principal component analysis. The d-dimensional arithmetic mean and dd come first.

O 2 i=1 I For the initial d-dimensional data collection, the covariance matrix S is calculated. After that, d eigen values are determined and Impacted by
If yi 1, then wTx + wO 1 - £i. are arranged with the value at the bottom. Assume the following eigenvalues (in
If yi is less than or equal to one, then wTx+wO—1+£i.

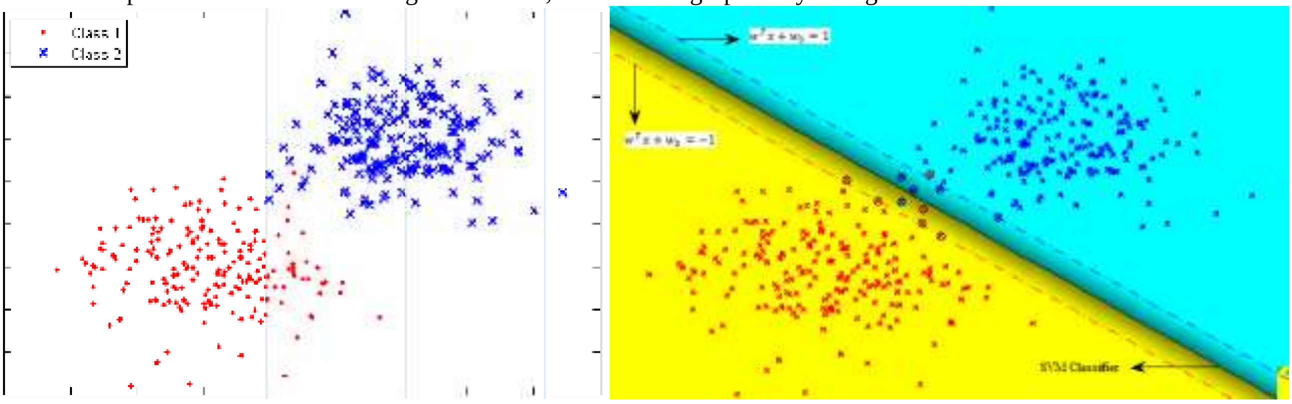
The corresponding eigenvalues (in decreasing order) are 1, 2,... d, while the eigenvectors (in the same format) are e1, e2,......ed. (any pair of eigenvectors, e1, e2,......, ed, is orthogonal to every other pair). After that, we have the first l eigen vectors.
and
I ≥ 0
The highest l eigenvalues, denoted by e1, e2,......el
In this case, we use a two-category system to organize our The numbers 2....l are used because they provide as a good starting point from which to extrapolate the d-.
The classifier equation is the first thing we study (epileptic seizure or not). information with l-dimensions in a spatial context. An acceptable value of l is determined by the magnitude of the difference between the lth and (l+1)th eigen values. Additional information and a

mathematical examination of PCA may be found in [12, 13].

C. Structural Feature Extraction

Originally, SVM was conceived as a means of managing the automated system's generalization capabilities. Ideally, a classifier's performance would be generalizable, or applicable to a wide variety of data types.

(the same as (1)) by using fifty percent of Ftr Mat's feature vectors as training data and resolving the optimization problem defined by (3) according to constraints defined by (4). Next, the remaining Ftr Mat feature vectors are classified using the separating hyper-plane. Learning the SVM classifier from mixed-class training data is shown graphically in Fig. 3.



(a) (b)

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

IV. EXPERIMENTAL RESULTS

In the work reported here, we use a pattern recognition strategy to the categorization of EEG signals. There were some EEG samples taken.

The precision is figured using:
TP+TN 100% = Accuracy
TP+FP+TN+FN
separated into sub-bands using 5-level discrete wavelet transform
It's a wavelet in the Daubechies format. There are 31 distinct The sub-bands of the details and the approximation were mined for statistical traits. All of the sample signals from sets A and E went through this same feature extraction method. After accumulating all of the feature vectors, a Ftr Mat with dimensions 31 by 200 was produced. Here, each column After categorization, data for accuracy, sensitivity, and specificity are summarized in Table II.

VI. CONCLUSION

Within the scope of this work, we present a pattern recognition strategy to is associated with a different characteristic, and each row.
Identify the epileptic attack. Methodology suggested here relies relates to a certain EEG sample. All of Ftr Mat's rows

were then scaled to a range of values between 0 and 1
means, power, standard deviation, kurtosis, skewness, entropy, and median are all extracted statistical properties.

x = xi-xmin, where I may take values from 1 to 200.

xmax–xmin
sub-bands extracted by DWT decomposition. The normalized values of are used to eliminate redundant characteristics in the observations.
The PCA algorithm is then used to further analyze the generated wavelet characteristics.
the value of the characteristic xi, where xi is an index On this specific row (the ith),
xmin= flattening of space-time dimensions. Dimensionality reduction in features by

equivalent to the lowest value in that row and xmax= the highest value in that row. The retrieved features' dimension was shrunk to 7.
Time and spatial complexity may both be reduced with PCA's assistance. The data is classified using SVM classification to determine if it represents epileptic seizures or other types of events. using principal component analysis. In the wake of dimension reduction
A party was held in our honor.
7×200The detection ratio and precision are both significantly improved.
100 out of 200 on the Ftr Mat features (50 from each class)
The results of using the suggested The application of a categorization scheme to were used as inputs for the SVM classifier's training process. We put the SVM classifier to the test using the remaining feature points.
Collection of electroencephalogram readings. The improved accuracy makes the technology an ideal aid for automatic categorization.
So that we can evaluate the classifier's efficiency
in regards to EEG readings.
test's level of sensitivity (True Positive Ratio- TPR) and specificity (True Development Given a certain hardware configuration and user
The Negative Ratio (TNR) values were determined.
using obfuscation
user-friendly design is potentially next-level stuff for the matrix. The formula used in the proposed study is described by Equations (6) and (7). insensitivity analysis and specificity utilizing the Real
Positive
(TP), (TN), (FP), and (FN) stand for "True Positive," "False Negative" (FN).

REFERENCES

Chambers, Saeid A., and Chambers. Analysis of Electroencephalogram (EEG) Data. In 2007, this book was published by John Wiley & Sons.

Pierre Gloor (1923–2003), Honored by A. Massimo. Publication information: Epilepsia, volume 45, issue 7, page 882, July 2004.
M. A. B. Brazier's book "Historically Speaking" is titled "The Brain's Electrifying Past."
Specifically, TNR equals TN plus FP at 100%.
The First Half-Century, BAR FRAME II, Macmillan, New York, 1961.
across numerous intracranial EEG electrode contacts: A report on four individuals," by M. D. Alessandro, R. Esteller, G. Vachtsevanos, A. Hinson, A. Echauz, and B.Litt. 2003, volume 50, issue 5 of the IEEE Transactions on Biomedical Engineering, pages 603-615.

Precision and Sensitivity Accuracy \s100.0% \s99.50% \s99.75%
Volume 85, Issue 3, Pages 261-281, 2003, Biosystems Engineering Time-frequency analysis in biosystem engineering, by B. P. Marchant.

An EEG signal is classified by A. Subasi using a combination of expert models and wavelet feature extraction. Article published in Expert Systems with Applications, volume 32, number 10, pages 1084-1093, 2007.
Inspired by "EEG signal classification using principal component analysis, independent component analysis, linear discriminant analysis, and support vector machines," by A. Subasi and M. Ismail Gursoy. [7]. Pages 8659–8666 of the 2010 edition of Expert Systems with Applications covered the subject in depth.
The study by A. Subasi, "Epileptic seizure detection using dynamic wavelet network," [8] is a good example of this kind of investigation. Expert Systems with Applications number 29, pages 343-355, 2005.
Title: "Classification of seizure based on the time-frequency picture of EEG data using HHT and SVM," Ku Fu, Ji Qu, Yu Chai, and Yu Dong. Biomedical Signal Processing and Control, Volume 13: 2014, Pages 15-22.
Dependence on recording region and mental state underlies the nonlinear, deterministic, and finite-dimensional patterns seen in time series of brain electrical activity. [10] Review of Physics Education 64(6):061907, 2001. R.G. Andrzejak, K. Lehnertz, F. Mormann, et al.
You should reference [11] H. Adeli, Z. Zhou, and N. Dadmehr. Article from 2003's Journal of Neuroscience Methods: "Wavelet transform analysis of electroencephalogram data from a patient with epilepsy," pages 69-87.
D. G. Stork, P. E. Hart, and R. O. Duda [12] John Wiley & Sons, 2001, Classifying Patterns, Second Edition.
K. Koutroumbas and S. Theodoridis [13]. Finding Variations. Elsevier/Academic Press brought out the 4th edition in 2009.
[14]
To learn more, read "An Overview of Support Vector Machines" by P. S. Sastry. J.C. Misra compiles recent advances in computing and data science. Narosa Publishing House of New Delhi, India, released a book in 2003.