(IJAER) 2013, Vol. No. 6, Issue No. III, September **e-ISSN: 2231-5152, p-ISSN: 2454-1796**

A METHOD BASED ON STATISTICAL WAVELET FEATURES, PCA, AND SVM FOR CLASSIFYING EEG SIGNALS

Muralidharan K

Lecturer in Electronics Engg. Government Polytechnic College,
Palakkad, Kerala-678551.

ABSTRACT

Electroencephalography is the study of the electrical signals generated by neural activity in the human brain. In this study, we suggest an automatic and effective method for classifying EEG signals. The five statistical features chosen for this paper are mean, standard deviation, energy, kurtosis, and skewness. Along with feature extraction, T-test comparison analysis of these features is also provided, taking into account both normal (non-PD) and people with PD symptoms' EEGs. Due to their superior generalization performance, support vector machines (SVMs) are a popular choice for classifying EEG signals. Due to the high processing complexity, SVM suffers. The EEG signal is divided into two categories using the suggested method: epileptic seizure or not. In the suggested method, we begin by decomposing the EEG signals into subbands and extracting the features using Discrete Wavelet Transform (DWT). Principal Component Analysis (PCA) uses these characteristics, which were extracted from the specifics and approximation coefficients of the DWT sub-bands, as input. The Support Vector Machine (SVM) and PCA are used to reduce the feature dimension and derive the support vectors, respectively, for the classification. The experimental procedures are carried out using real and typical datasets. Given that EEG has a pattern that is similar to spike-waves and that the Daubechies sequence has higher scalability and flexibility for weighting boundary difficulties, the Daubechies-2 wavelet function is appropriate for EEG [1]. Four-level decomposition, on the other hand, is used because it captures the EEG's major frequency bands.

Keywords: *Discrete Wavelet Transform; Electroencephalogram; Pattern Recognition; Principal Component Analysis; Support Vector Machine.*

INTRODUCTION

The examination of the electrical impulses generated by the brain is called an electroencephalogram. The human brain produces electrical signals that are a representation of mental and physical state. The in-depth examination of those EEG signals is beneficial in the detection and classification of event-related potentials, the detection and prediction of seizures, brain-computer interface research, the study of mental disorders such as psychiatric disorders, dementedness, and sleep signal analysis. The EEG signal waves are further separated into five key sub-bands supplemented by frequency ranges for a deeper comprehension of human behavior.

(IJAER) 2013, Vol. No. 6, Issue No. III, September **e-ISSN: 2231-5152, p-ISSN: 2454-1796**

EEGs are pictures or recordings of the electrical potentials that the brain generates. Since Hans Berger began recording rhythmic electrical activity from the human scalp, analysis of EEG activity has primarily been done in clinical settings to pinpoint diseases and epilepsies. In the past, neurophysiologists, who are qualified to qualitatively distinguish between normal EEG activity and anomalies found in EEG recordings, could only interpret the EEG by visual inspection. A variety of ways to assess EEG changes are now potentially applicable because to advancements in computer technology.

The brain is one of the most significant body parts. The brain is the vertebrate central nervous system, and it is made up of both gray and white matter in addition to being contained within the skull and spinal cord. The brain is the primary center of regulation and manages all bodily functions, including breathing and heartbeat. It is a very intricate system that displays vibrant spatiotemporal dynamics throughout the entire body. And it frequently experiences some kind of issue or malfunction. Epilepsy is one of the most prevalent mental illnesses. Today, this disorder affects around 60 million people. The electroencephalogram, which is curtailed as EEG, is clinically used to research cerebrum problems or to analyze different mental functionalities. The main effort to gauge this cerebrum problem movement was finished by English doctor Richard Caton in 1875. The electroencephalographic record is one of the main apparatuses for the investigation of cerebrum electrical movement and for the finding of neurological illnesses. It is likewise the electrical translation of the heart action. Epileptic seizures are fundamentally the appearance of epilepsy. Electroencephalograph (EEG) records can provide important knowledge and work on comprehension of the components causing epileptic issues through cautious investigation. Epileptic seizures are caused by the transitory electrical aggravation of the mind. Epileptic seizures at times slip by without everyone's notice, and now and again, chomped mistook for different occasions, winning based on their introductions, for example, strokes, falls, and headaches. Out of each and every 100 people, one is encountering a seizure at some point in for what seems like forever. Sadly, the event of epileptic seizure appears to be some way or another erratic and very little understood.5 In the conclusion of epilepsy, the recognition of epileptiform releases in the EEG is a significant part. EEG signals are non-fixed. The main, valuable and practical methodology for the investigation of epilepsy is EEG. Likewise, EEG is a graphical portrayal of cardiovascular action which involves essential measure for the ID of different heart sicknesses and heart irregularities. EEG signal is fundamentally a one-layered natural sign. Investigating of EEG signal depends on its recurrence content. Consequently we can say that translation of EEG signal depends on the force of the frequencies it contains.

(IJAER) 2013, Vol. No. 6, Issue No. III, September **e-ISSN: 2231-5152, p-ISSN: 2454-1796**

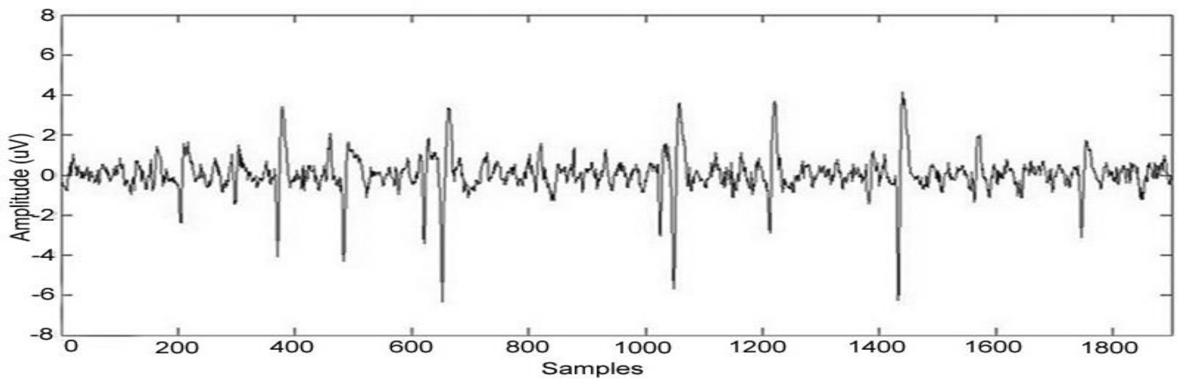


Figure 1: 1D EEG Signal

RELATED WORK

The electrical signs for mind movement were first kept by the English researcher Richard Caton in 1875. Hans Berger began the investigation of EEGs from the human cerebrum in 1920 [2]. Epilepsy is a Greek word, and that means 'to seize or assault'. The exceptionally essential ideas of epilepsy can be tracked down in old Indian medication (4500–1500 BC) as apasmara, and that signifies "loss of awareness". Babylonian tablet in the English Exhibition hall in London additionally gives the point by point information about the epileptic sickness and its fix [1]. Kaufman connected the epileptic assaults with strange electrical releases [3]. A large portion of the epilepsy examination techniques created in the twentieth century depended on the idea of a visual review of the EEG and exceptionally talented electroencephalographies. Notwithstanding, with the progression in the field of sign handling and example acknowledgment, different programmed strategies of epileptic seizure location have been created in most recent twenty years [6], [9]. Phantom examination based include extraction strategy gives unfortunate outcomes to EEG arrangement as the recurrence area data is given at the expense of time space data, for example, the adequacy conveyance and EEG design. Consequently, both time and recurrence space based highlight extraction calculations, for example, Discrete Wavelet Change (DWT) are being utilized in momentum research [4],[5]. The other benefit of DWT over ghastly examination is its appropriateness for investigation of non-fixed signals like EEG [6] - [8]. Hilbert-Huang Changed (HHT) based approach as of late distributed by [9].

FEATURE SELECTION

The selection of features is a critical phase in categorization. To achieve the greatest results, it is best to feed the chosen features. The system may get overloaded due to redundant features, which could prevent the best outcome from happening. Therefore, cutting down on the number of features will help the classifier learn more effectively and perform well.[6] Less complexity results from removing unnecessary features and learning solely from relevant features. In this research, the 8 pertinent features are identified from the 16 dimension features using the genetic algorithm (GA) technique [7].

CLASSIFICATION IN MACHINE LEARNING

One of the key methods used by machine learning to classify the categorical labels in data analysis is classification. The procedure can be divided into two phases. The database tuples and their associated class labels were used to educate the system during the learning stage. The system is tested with test data for categorization in the second stage. The percentage of test data that the classifier correctly categorizes serves as a measure of the system's performance.[8]

K-fold cross-validation

Data must be divided into subsets in order to validate the complete input. The statistical technique known as "k-fold cross-validation" divides the sample data into k subgroups. K-1 subsets of the total number of subsets are used to train the proposed system, and the final subset is utilized to test its performance. In order to use all of the subsets as validation data, this method will be done k times (folds) in total. To measure the estimation accuracy, the data from the k-folds can be averaged. This method's goal is to validate all of the samples.[6] The 10-fold cross-validation approach is employed in this study's studies to verify the authenticity of the provided samples.

Support vector machine

In order to categorize normal and abnormal features, Support Vector Machine (SVM) constructs a hyperplane. Due to SVM classifier's excellent accuracy, it can handle large dimensional data.[9] An algorithm called Kernel can create nonlinear decision boundaries.[10] SVM functions essentially as a linear two-class classifier. The designations "+1" and "1" classify the characteristics of the two classes.[2-3] The SVM classifier is utilized in this study to categorize EEG signals.

Wavelet Transform

The electrical exercises of the cerebrum since the 1930s have been estimated by utilizing surface anodes associated with the scalp. However, these days, different numerical apparatuses, for example, Fourier Change (FT), Quick Fourier Change (FFT), Brief Time Frame Fourier Change, and Wavelet Change (WT), have been presented for EEG signal element extraction. Anyway, in Quick Fourier Change, there was data misfortune at no time like in the present space and just ghastly data in the recurrence area. To overcome the issues connected with FFT, a STFT was presented that addressed the sign in both time and recurrence areas, utilizing moving window capability. The principal issue related to STFT is that it doesn't give multi-goal data on the signs as it generally has a steady size. To overcome the issues connected with Fourier change, Fat Fourier change, and Brief time frame Fourier change, a strong technique was proposed in the late 1980s, known as wavelet change. Wavelet change can be thought of as an expansion of Fourier change, and furthermore, as opposed to dealing with a solitary scale (time or recurrence), it chips away at a multi-scale premise and furthermore resolves the issues connected with non-fixed signals. The wavelet change hypothesis has tracked down many fascinating and intriguing applications with regards to the field of computerized signal handling. As of late, wavelet examination has assumed a significant role in investigating time-space signals. A wavelet is a sort of time-recurrence investigation that gives data about both recurrence and time inside signals. Investigation by wavelet addresses a unique kind of direct change of signs and, furthermore, actual information addressed by the signs about processes and the actual properties of mediums and items. Wavelet Change has been more proficient for signal examination in contrast with other change strategies, for example, Fourier change and a brief time frame Fourier Change. The primary benefit of wavelet change is that it has a fluctuating window size that is wide at low frequencies and thin at high frequencies, prompting an ideal time recurrence goal in all recurrence ranges (for example, it holds multiresolution properties). In math, a wavelet series is a portrayal of square integrable, which can be genuine or complex. A wavelet is a wave like swaying with sufficiency that starts at nothing, increments, and afterward diminishes back to nothing. Wavelets are for the most part created to have explicit properties that make them helpful for signal handling. Wavelets can be consolidated utilizing an opposite, shift, increase and incorporate strategy known as convolution.13 Wavelets changes are extensively partitioned into three classes: persistent, discrete, and multiresolution-based.

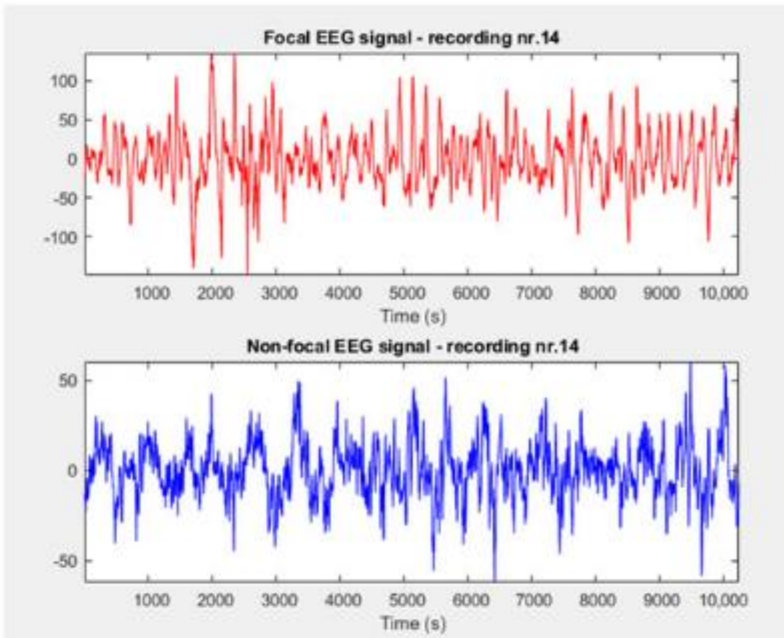
Support Vector Machine (SVM)

SVM is another sort of classifier that is propelled by or in view of two ideas. The first idea incorporates the changing of information into a high-layered space. This idea can change complex issues (with complex choice surfaces) into more straightforward issues that utilize direct discriminant capabilities. What's more, the second idea of SVMs is propelled by preparing and utilizing just those data sources that are close to the choice surface as they give the main data about the grouping. The SVM calculation depends on the measurable learning hypothesis. SVM is utilized for information grouping, design acknowledgment, bioinformatics applications, and furthermore for relapse examination as a result of its exactness and capacity to manage a huge number of indicators. The help vector classifier enjoys many benefits on its own. In SVM, nonlinear limits can be utilized without requiring a lot of extra computational exertion. In addition, the execution of SVM is extremely aggressive with different techniques. A downside of SVM is the issue of intricacy, which isn't of the request for the component of the examples but of the request for the quantity of tests.

EXPERIMENTAL RESULT AND DISCUSSION

We first retrieved the EEG signal's features before expressing the classification findings in terms of criteria including sensitivity, specificity, and accuracy in the current study. Wavelet transform was used for the feature extraction. For wavelet, the discrete wavelet transform is used to extract features, and the wavelet type utilized is the Daubechies wavelet, or db3 (level 3). To reduce the signal's dimension, the EEG signal is first divided into wavelet coefficients using the wavelet decomposition scheme and db3 (level 3). The features that describe the signal are based on these decomposed coefficients. Classification is completed when the EEG signal's features have been extracted.

Artificial neural networks (ANN) and support vector machines (SVM) are employed in this study as classifiers for both normal and epileptic patients. The classifiers are fed the data that was obtained after feature extraction as input.



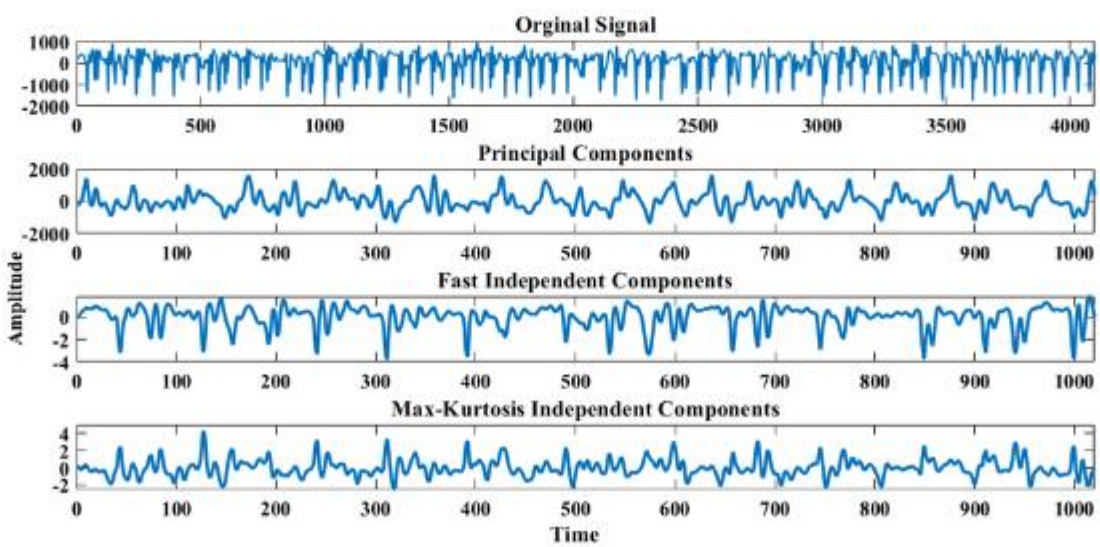


Figure 2: First components of PCA, FICA and KICA from a given EEG signal.

Classification using Support Vector Machine (SVM)

The training dataset for SVM-based classification uses the kernel, a Gaussian radial basis function. Training and testing follow an 80:20 ratio. The recognition rate is ultimately calculated using a confusion matrix.

The performance analysis of the classifier is tested using parameters like Sensitivity, also known as the true positive ratio (TPR), Specificity, also known as the true negative ratio (TNR), and Accuracy.

CONCLUSION

The wavelet transform is suitable for locating discontinuities or singularities where high-frequency components predominate because of its nature, which makes it well suited to the analysis of signals where a more accurate time resolution is required for higher frequencies than for lower ones. We have suggested a wavelet-based feature extraction method for categorizing epileptic EEG signals. Careful EEG analysis offers insightful knowledge about how the brain works and can be helpful in identifying brain disorders, particularly in the case of epilepsy. Insuring that epileptic patients receive quality care is extremely crucial. Supportive systems will enable professionals to make more accurate diagnoses. An automated seizure detection model had been successfully introduced in this paper. By employing DWT at level 4, the EEG signals were divided into sub-bands using the db2 wavelet. The research on the proposed strategy is clear from the other approaches to pattern recognition that the researchers looked into for various combinations of data sets.

REFERENCES

[1]. Acharya, U. Rajendra, et al. "Use of principal component analysis for automatic classification of epileptic EEG activities in wavelet framework." Expert Systems with Applications 39.10 (2012): 9072-9078.

[2]. Mahajan, Kavita, M. R. Vargantwar, and Sangita M. Rajput. "Classification of EEG using PCA, ICA and Neural Network." International Journal of Engineering and Advanced Technology 1.1 (2011): 80-83.

[3]. Hsu, Wei-Yen, et al. "Wavelet-based envelope features with automatic EOG artifact removal: Application to single-trial EEG data." Expert Systems with Applications 39.3 (2012): 2743-2749.

[4]. Gandhi, Tapan, Bijay Ketan Panigrahi, and Sneh Anand. "A comparative study of wavelet families for EEG signal classification." Neurocomputing 74.17 (2011): 3051-3057.

[5]. Uğuz, Harun. "A hybrid system based on information gain and principal component analysis for the classification of transcranial Doppler signals." Computer methods and programs in biomedicine 107.3 (2012): 598-609.

[6]. Liu, Yinxia, et al. "Automatic seizure detection using wavelet transform and SVM in long-term intracranial EEG." IEEE transactions on neural systems and rehabilitation engineering 20.6 (2012): 749-755.

[7]. Acharya, U. Rajendra, et al. "Automated diagnosis of epileptic EEG using entropies." Biomedical Signal Processing and Control 7.4 (2012): 401-408.

[8]. Yohanes, Rendi EJ, Wee Ser, and Guang-bin Huang. "Discrete wavelet transform coefficients for emotion recognition from EEG signals." 2012 annual international conference of the IEEE engineering in medicine and biology society. IEEE, 2012.

[9]. Du, Xian, et al. "Classification of epilepsy using high-order spectra features and principle component analysis." Journal of medical systems 36 (2012): 1731-1743.

[10]. Acharya, U. Rajendra, et al. "Application of non-linear and wavelet based features for the automated identification of epileptic EEG signals." International journal of neural systems 22.02 (2012): 1250002.