Multi-class SVM for EEG Signal Classification Using Wavelet Based Approximate Entropy

A.S. Muthanantha Murugavel* and S. Ramakrishnan

Department of Information Technology,
Dr. Mahalingam College of Engineering and Technology,
Pollachi – 642 003, Tamilnadu, India
{as_ananth2k1,ram_f77}@yahoo.com

Abstract. In this paper, we have proposed a novel wavelet based approximate entropy for feature extraction and a novel Multi-Class Support Vector Machine (MSVM) for the multi-class electroencephalogram (EEG) signals classification with the emphasis on epileptic seizure detection. The aim was to determine an effective classifier and features for this problem. Wavelets have played an important role in biomedical signal processing for its ability to capture localized spatial-frequency information of EEG signals. The MSVM works well for high dimensional, multi-class data streams. Decision making was performed in two stages: feature extraction by computing the wavelet based approximate entropy and classification using the classifiers trained on the extracted features. We have compared the MSVM with Probabilistic Neural Network (PNN) by evaluating with the benchmark EEG dataset. Our experimental results show that the MSVM with wavelet based approximate entropy features gives high classification accuracies than the existing classifier.

Keywords: Electroencephalogram (EEG) signals classification, epileptic seizure detection, Multi-Class Support Vector Machine, Wavelet Transform, Approximate Entropy, Probabilistic Neural Network (PNN).

1 Introduction

The electroencephalogram (EEG) is a complex and aperiodic time series, which is a sum over a very large number of neuronal membrane potentials. Despite rapid advances of neuroimaging techniques, EEG recordings continue to play an important role in both the diagnosis of neurological diseases and understanding the psychophysiological processes. In order to extract relevant information from recordings of brain electrical activity, a variety of computerized-analysis methods have been developed. Most methods are based on the assumption that the EEG is generated by a highly complex linear system, resulting in characteristic signal features like nonstationary or unpredictability [1]. Much research with nonlinear methods revealed that the EEG is generated by a chaotic neural process of low dimension [2]–[4]. As in traditional pattern recognition systems, classification of biomedical

* Corresponding author.

N. Meghanathan et al. (Eds.): CCSIT 2012, Part II, LNICST 85, pp. 335–344, 2012.

signals consists of two main modules namely Feature Extraction and Feature Classification. In recent papers, for studying and analyzing the behavior of EEG signals, chaos theory was used [1]–[4]. To quantify the complexity of EEG signals, numbers of entropy estimators are available. This proposed technique uses approximate entropy (ApEn) as the input feature. The general structure of developed EEG signals classification model has two modules (Fig. 1). A significant contribution of our work was the composition of composite features, which were used to train novel classifier. PNN is a type of radial basis network. For the classification and comparison we used the benchmarked datasets (EEG Signals) which includes five classes. The paper is organized as follows. In Section 2, we briefly presented the literature survey that has been performed. In Section 3 we described about the benchmark dataset, our proposed methodology with feature extraction method and classification techniques that are considered. In Section 4, we compared the results of the proposed classifiers using the features with other existing classifier. And Section 5 concludes the paper.

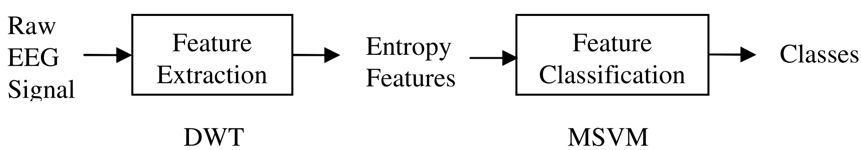

Fig. 1. General Structure of developed EEG- signals classification

2 Literature Survey

Automatic analysis and diagnosis of epilepsy based on EEG recordings is started in the early 1970s. Today, computer-based analysis addresses two major problems: 1) interictal event detection 2) epileptic seizure analysis [1]. Various feature extraction techniques have been used for the classification of EEG signals with the emphasis on seizure detection. Non-linear based feature extraction technique uses Correlation Dimension, Lyapunov Exponent and Standard Deviation for extracting the features of EEG signals [1][2]. Entropy is a term of thermodynamics that is used to describe amount of disorder in a system.Entropy based technique uses Approximate Entropy (ApEn) as the input feature [4] [6]. Wavelet based technique uses Max, Min, Mean and Standard Deviation [2] [7] [8] [9]. Time Frequency based technique, Feature extraction based on Local Variance [3]. The benchmarked dataset that have been used in the existing works was used to compare the proposed method with the existing methods. The EEG signal classification techniques are divided into three broader categories: Conventional classifiers such as Linear Discriminant Analysis [10], Support Vector Machine [8], and Naïve Bayes [1]. Neural Networks such as MLPNN [8], PNN [10], and RBFNN [6], Combinational classifiers such as Boosting, Voting and Stacking [11].

3 Proposed Methodology

3.1 Dataset Description

We have used the publicly available benchmark EEG dataset described in Andrzejak et al. [8]. The complete data set consists of five sets (denoted A–E) each containing 100 single-channel EEG segments. Sets A and B consisted of segments taken from surface EEG recordings that were carried out on five healthy volunteers using a standardised electrode placement scheme. Volunteers were relaxed in an awake-state with eyes open (A) and eyes closed (B), respectively. Sets C, D, and E originated from EEG archive of pre-surgical diagnosis. EEGs from five patients were selected, all of whom had achieved complete seizure control after resection of one of the hippocampal formations, which was therefore correctly diagnosed to be the epileptogenic zone. Segments in set D were recorded from within the epileptogenic zone, and those in set C from the hippocampal formation of the opposite hemisphere of the brain. While sets C and D contained only activity measured during seizure free intervals, set E only contained seizure activity.

3.2 Extraction of Features

Generally feature extraction is transforming the raw input data into set of features. Wavelet Transform (WT) is a spectral estimation technique in which any general function can be expressed as an infinite series of wavelets. Abnormalities in the EEG in serious psychiatric disorders are many times too subtle to be detected using conventional techniques, such as Fourier transform. WT is specific appropriate for analysis of non-stationary signals. It is well suited for locating transient events, which always occur during epileptic seizure. The decomposition of the signal leads to a set of coefficients called wavelet coefficients. The key feature of wavelets is the time-frequency localization.

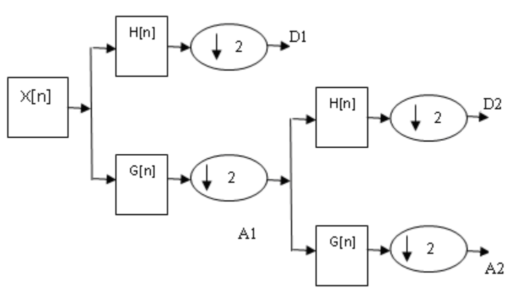

Fig. 2. Wavelet Decomposition of EEG signal into two levels of sub-bands

The decomposition of the signal into the different frequency bands is merely obtained by consecutive high-pass and low-pass filtering of the time domain signal. The procedure of multi-resolution decomposition of a signal x[n] is schematically shown in Fig. 2. Each stage of this scheme consists of two digital filters and two down-samplers by 2. The first filter, h[n] is the discrete mother wavelet, high pass in nature, and the second, g[n] is its mirror version, low-pass in nature. The detail at level j is defined as

$$D_j = \sum_{k \in Z} a_{\mathrm{j,\,k}} \, \psi_{j,\,k}(t) \tag{1}$$

The approximation at level J is defined as

$$A_j = \sum_{j>J} D_j \tag{2}$$

It becomes obvious that

$$A_{j-1} = A_j + D_j \tag{3}$$

and

$$f(t) = A_j + \sum_{j \le J} D_j \tag{4}$$

The down-sampled outputs of first high-pass and low-pass filters provide the detail, D1 and the approximation, A1, respectively.

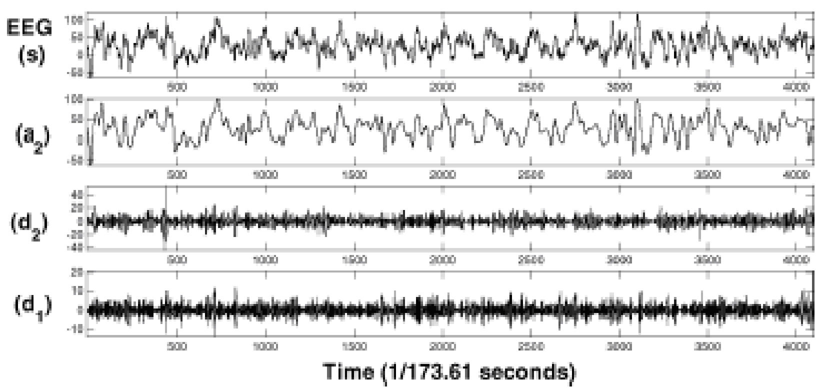

Fig. 3. Level 2 decomposition of the band-limited EEG into three EEG sub bands using fourth-order Daubechies wavelet (s = a2+d2+d1)

Wavelet has several advantages, which can simultaneously possess compact support, orthogonality, symmetry, and short support, and high order approximation. We experimentally found that time-frequency domain feature provides superior performance over time domain feature in the detection of epileptic EEG signals. Usually, tests are performed with different types of wavelets and the one, which gives maximum efficiency, is selected for the particular application. The smoothing feature of the Daubechies wavelet of order 4 (db4) made it more appropriate to detect changes of EEG signals. The EEG sub bands of a2, d2 and d1are shown in fig. 3.

Wavelet based Approximate Entropy (ApEn)

The proposed system makes use of a single feature called ApEn for the epileptic detection. The ApEn is a wavelet-domain feature that is capable of classifying complex systems. The value of the ApEn is determined as shown in the following steps. Table 1 represents the extracted wavelet based Approximate Entropy (ApEn) features for the decomposed sub bands.

Table 1. Wavelet based ApEn values used for Training and testing the classifier

Data Points	Wavelet based ApEn values used for training		Wavelet ApEn values used for testing	
	Intracranial	Normal	Intracranial	Normal
173	1380	1380	920	920
256	932	932	620	620
512	466	466	310	310
1024	233	233	155	155
2048	116	166	77	77

1) Let the data sequence containing N data points be X = [x(1), x(2), x(3), . . . , x(N)].

2) Let x(i) be a subsequence of X such that x(i) =[x(i), x(i + 1), x(i + 2), . . . , x(i + m − 1)] for $1 \leq i \leq N - m$, where m represents the number of samples used for the prediction.

3) Let r represent the noise filter level that is defined as

$$r = k \times SD \tag{5}$$

for k = 0, 0.1, 0.2, 0.3, . . . , 0.9
where SD is the standard deviation of the data sequence X.

4) Let {x(j)} represent a set of subsequences obtained from x(j) by varying j from 1 to N. Each sequence x(j) in the set of {x(j)} is compared with x(i) and, in this process, two parameters, namely Cim(r) and Cim+1(r) are defined as follows:

$$Cim(r) = N\text{-}m \tag{6}$$

where k = 1, if $|x(i) - x(j)| \leq r$ for $1 \leq j \leq N - m$
0, otherwise

and

$$\text{Cim+1(r) =N-m} \tag{7}$$

with conditions depicted by (A) as shown at the bottom of the page.

5) We define Φm(r) and Φm+1(r) as follows:

$$\Phi\text{m(r) = N-m} \tag{8}$$

$$\Phi\text{m+1(r) = N-m} \tag{9}$$

Small values of ApEn imply strong regularity in a data sequence and large values imply substantial fluctuations [11].

3.3 Classifiers Used for Classification

Probabilistic Neural Network (PNN)

The PNN was first proposed by Specht [14]. Various multi-class problems can be handled by a single PNN.

$$\phi_{ij}(x) = \frac{1}{(2\pi)^{d/2}\sigma^d}\exp\left[-\frac{(x-x_{ij})^T(x-x_{ij})}{2\sigma^2}\right] \tag{10}$$

where d denotes the dimension of the pattern vector x, is the smoothing parameter, and x_{ij} is the neuron vector [8].

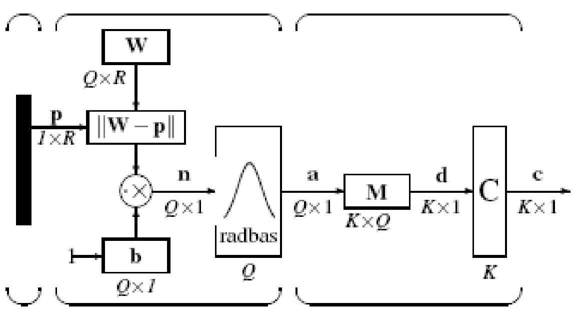

Fig. 4. PNN Architecture, R: number of features, Q: number of training samples, K: number of classes

On receiving a pattern x from the input layer, the neuron x_{ij} of the pattern layer computes its output. The structure of PNN is depicted in Fig. 4.

Multi-class Support Vector Machine (MSVM)

The SVM is basically a binary classifier that can be extended by modifying several of its kind into a multi-class classifier. The general structure of SVM is depicted in Fig. 5. The SVM is a one pass incremental algorithm that does not require the following such as a sliding window on the data stream and monitoring the performance of the classifier as data points are streaming. The principle idea is to assign a binary code word of length N, denoted here

$$t(c)\in\{-1,0,1\}^N \tag{11}$$

$$t(c)\in\{0,1\}^N, t(c)\in\{-1,1\}^N \tag{12}$$

or even for each class c. The result is the code matrix T represented here

$$T = \begin{pmatrix} t(1) \\ t(2) \\ . \\ . \\ t(N) \end{pmatrix} = \begin{pmatrix} t_{11}t_{12}\ldots.t_{1N} \\ t_{21}t_{22}\ldots.t_{2N} \\ . \\ . \\ t_{N1}t_{N2}\ldots t_{NN} \end{pmatrix} \tag{13}$$

Now each column defines a separation of the classes in two subsets f¡1; 1g, 0 valued elements are simply ignored. Each column is fed into a separate classifier for learning and recognition. The result is another codeword tL which can be compared with the existing N code words by using Hamming or other distance measures. For dichotomies, a soft margin classifier can be defined. It can be understood as mapping the property $y_n f(\overrightarrow{x_n}) \geq \rho$ with $(\overrightarrow{x_n}, y_n)$ {training set} and some positive constant giving the margin. The function f is also called the embedding. To avoid over fitting some slack variables are also introduced

$$\text{Maximize} \quad \frac{\lambda}{\mathrm{M}}\sum_{m=1}^{M} \varepsilon_m + \Omega\{f\} \text{ (the margin)} \tag{14}$$

This margin maximization alone is nothing new, but the notation can be extended to polychotomial problems. The minimal relative difference in distance between f, the correct target t(y) and any other target t(c). The new optimization problem can now be written in the following way.

$$\text{Minimize } \frac{\lambda}{\mathrm{M}}\sum\nolimits_{m=1}^{\mathrm{M}} \varepsilon_m + \Omega\{f\} \tag{15}$$

This optimization problem is a multi-class classifier using the distance measure function d and large soft margins.

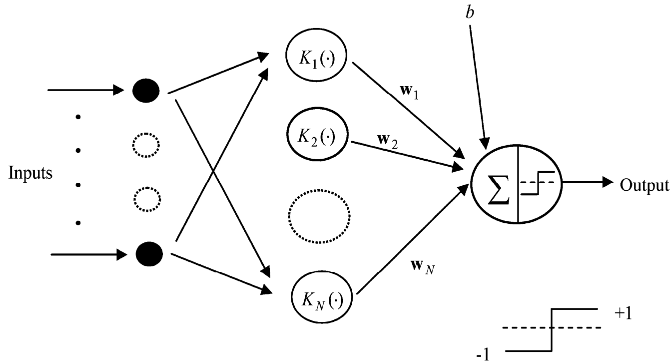

Fig. 5. General Architecture of SVM

4 Results and Discussion

In our work, we employed a discrete wavelet transform for extracting approximate entropy features from the dataset in order to extract temporal information. The wavelet base ApEn possesses good characteristics such as robustness in the characterization of the epileptic patterns and low computational burden. n. ApEn values are computed for selected combinations of m, r, and N. The values of m, r, and N that are used for the experiments are as follows: m = 1, 2, 3; r = 0%–90% of SD of the data sequence in increments of 10%; and N = 4097. Wavelet based ApEn values are computed for both normal and epileptic EEG signals and are fed as inputs to the two neural networks. The potentiality of the wavelet based ApEn to discriminate the two signals, namely, normal and epileptic EEG signals depends on the values of m, r, and N.

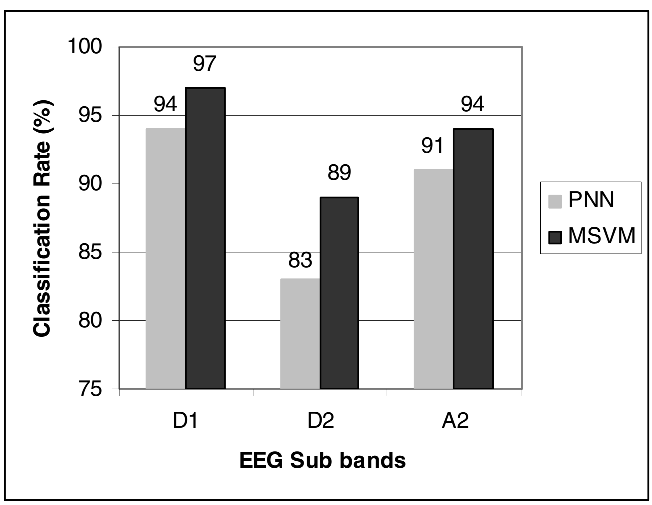

Fig. 6. Comparison of classifiers such as Probabilistic Neural Network and Multi-class Support Vector Machine based on classification rate

For the classification of EEG signals we have used a novel multi-class SVM. Among the available 100 EEG data sets, 50 data sets are used for training and the remaining data sets are used for testing the performance of the neural network and SVM classifiers. Computational cost and the classification time for the proposed MSVM classifier depend on the number of support vectors required for the design of the classifier and the kernel employed. Increase in number of support vectors lead to increase in computational requirements. It is shown that our wavelet based ApEn possesses good characteristics such as robustness in the characterization of the epileptic patterns and low computational burden. The performance in the rate of classification of the two classifiers such as PNN and multi-class SVM have been compared with different sub bands and which is shown in Fig. 6. The Multi-class SVM using wavelet based Approximate Entropy features for the sub-band D1 gives superior performance in terms of classification rate.

5 Conclusion

The MSVM has shown great performance since it measures the predictability of the current amplitude values of a physiological signal based on its previous amplitude values. A robust and computationally low-intensive feature, wavelet based ApEn has been used as the feature for the proposed system. Besides this, the PNN provided encouraging results, which could have originated from the architecture of the PNN. The performance of the other neural network was not as high as the MSVM. The results of the present paper demonstrated that the MSVM with wavelet based Approximate Entropy feature can be used in the classification of the EEG signals by taking the misclassification rates into consideration. In current work, focus was put on normal and epileptic EEG signal classification. In the next stage of research, the results from this preliminary study will be expanded to include a more complete range of datasets.

Acknowledgment. The authors wish to thank Andrzejak et al., 2001 for the benchmark EEG dataset available:(http://www.meb.unibonn.de/epileptologie/science/physik/eegdata.html).

References

1. Adeli, H., Ghosh-Dastidar, S., Dadmehr, N.: A Wavelet-Chaos Methodology for Analysis of EEGs and EEG Subbands to Detect Seizure and Epilepsy. IEEE Transactions on Biomedical Engineering 54(2) (2007)
2. Ghosh-Dastidar, S., Adeli, H., Dadmehr, N.: Mixed-Band Wavelet-Chaos- Neural Network Methodology for Epilepsy and Epileptic Seizure Detection. IEEE Transactions on Biomedical Engineering 54(9) (2007)
3. Ghosh-Dastidar, S., Adeli, H., Dadmehr, N.: Principal Component Analysis-Enhanced Cosine Radial Basis Function Neural Network for Robust Epilepsy and Seizure Detection. IEEE Transactions on Biomedical Engineering 55(2) (2008)

4. Srinivasan, V., Eswaran, C., Sriraam, N.: Approximate Entropy-Based Epileptic EEG Detection Using Artificial Neural Networks. IEEE Transactions on Information Technology in Biomedicine 11(3) (2007)
5. Tzallas, A.T., Tsipouras, M.G., Fotiadis, D.I.: Epileptic Seizure Detection in EEGs Using Time–Frequency Analysis. IEEE Transactions on Information Technology in Biomedicine 13(5) (2009)
6. Jahankhani, P., Kodogiannis, V., Revett, K.: EEG Signal Classification Using Wavelet Feature Extraction and Neural Networks. In: IEEE John Vincent Atanasoff 2006 International Symposium on Modern Computing, JVA 2006 (2006), 0-7695-2643-8/06
7. Muthanantha Murugavel, A.S., Ramakrishnan, S.: Classification of EEG Signals Using Weighted Probabilistic Neural Network and Multiwavelet Transformation. In: Proceedings of the International Conference on Operation Research Applications in Engineering and Management (ICOREM), May 27-29 (2009)
8. Firpi, H., Goodman, E.D., Echauz, J.: Epileptic Seizure Detection Using Genetically Programmed Artificial Features. IEEE Transactions on Biomedical Engineering 54(2) (2007)
9. Güler, I., Übeyli, E.D.: Multiclass Support Vector Machines for EEG-Signals Classification. IEEE Transactions on Information Technology In Biomedicine 11(2) (2007)
10. Garrett, D., Peterson, D.A., Anderson, C.W.: Comparison of Linear, Nonlinear, and Feature Selection Methods for EEG Signal Classification. IEEE Transactions on Neural Systems and Rehabilitation Engineering 11(2) (2003)
11. Hazarika, N., Tsoi, A.C., Sergejew, A.A.: Nonlinear Considerations in EEG Signal Classification. IEEE Transactions on Signal Processing 45(4) (1997)
12. Haselsteiner, E., Pfurtscheller, G.: Using Time-Dependent Neural Networks for EEG Classification. IEEE Transactions on Rehabilitation Engineering 8(4) (2000)
13. Townsend, G., Graimann, B., Pfurtscheller, G.: Continuous EEG Classification During Motor Imagery—Simulation of an Asynchronous BCI. IEEE Transactions on Neural Systems and Rehabilitation Engineering 12(2) (2004)
14. Suffczynski, P., Lopes da Silva, F.H., Parra, J., Velis, D.N., Bouwman, B.M., van Rijn, C.M., van Hese, P., Boon, P., Khosravani, H., Derchansky, M., Carlen, P., Kalitzin, S.: Dynamics of Epileptic Phenomena Determined From Statistics of Ictal Transitions. IEEE Transactions on Biomedical Engineering 53(3) (2006)