ISSN:0975-1459
Journal of Pharmaceutical Sciences and Research
www.jpsr.pharmainfo.in

Principal Component Analysis as a Dimensionality Reduction Technique and Sparse Representation Classifier as a Post Classifier for the Classification of Epilepsy Risk Levels from EEG Signals

R.Harikumar[1], P.Sunil Kumar[2]
[1]Professor, Department of ECE, Bannari Amman Institute of Technology, India
[2] Research Scholar, Department of ECE, Bannari Amman Institute of Technology, India

Abstract

The main aim of this paper is to perform the analysis of Principal Component Analysis (PCA) as a Dimensionality Reduction technique and Sparse Representation Classifier (SRC) as a Post Classifier for the Classification of Epilepsy Risk levels from Electroencephalography signals. The data acquisition of the EEG signals is performed initially. Then PCA is applied here as a dimensionality reduction technique and then Sparse Representation Classifier is used for the Classification of Epilepsy Risk levels from EEG signals. The performance of the PCA with the SRC are compared based on the parameters such as Performance Index (PI) and Quality Value (QV).

Keywords

EEG Signals, SVD, Sparse, Performance Index, Quality Values

1. Introduction

Epilepsy is a collection of neurological disorders which is mainly characterized by epileptic seizures [1]. They can vary from short to long periods of incessant and vigorous shaking. The seizures always tend to occur in epilepsy and the underlying cause is not known here [2]. The exact reason for the occurrence of epilepsy is not known, but factors such as brain injury and stroke, brain-tumour, genetic mutations have contributed a lot to it [3]. The electroencephalogram is used to confirm the presence of epilepsy but just with a normal test, the condition cannot be ruled out easily [4]. Some seizures are controllable with the help of medications and in conditions where a seizure does not respond to medication, appropriate life-style changes have to be made along with neuro-stimulation and surgery [5]. Thus epilepsy is always characterized by a long term risk of recurrent seizures. The factors such as the involvement of a particular part of a brain and the person's age are considered for a seizure development.

2. Materials And Methods

2.1 Data Acquisition of EEG Signals

For the performance analysis of the epilepsy risk levels using singular value decomposition as a dimensionality reduction technique and sparse representation classifier as a post classifier, the raw EEG data of 20 epileptic patients who were under treatment in the Neurology Department of Sri Ramakrishna Hospital, Coimbatore in European Data Format (EDF) are taken for study. The preprocessing stage of the EEG signals is given more attention because it is vital to use the best available technique in literature to extract all the useful information embedded in the non-stationary biomedical signals. The EEG records which were obtained were continuous for about 30 seconds and each of them was divided into epochs of two second duration. Generally a two second epoch is long enough to avoid unnecessary redundancy in the signal and it is long enough to detect any significant changes in activity and to detect the presence of artifacts in the signal. For each and every patient, the total number of channels is 16 and it is over three epochs. The frequency is considered to be 50 Hz and the sampling frequency is considered to be about 200 Hz. Each and every sample corresponds to the instantaneous amplitude values of the signal which totals to 400 values for an epoch. The total number of artifacts present in the data is four. Chewing artifact, motion artifact, eye blink and electromyography (EMG) are the four number of artifacts present and approximately the percentage of data which are artifacts is 1%. No attempts were made to select certain number of artifacts which are of more specific nature. The main objective to include artifacts is to differentiate the spike categories of waveforms from non spike categories.

3. Principal Component Analysis

All suboptimal transforms such as the DFT and DCT decompose the signals into a set of coefficients, which do not necessarily represent the constituent components of the signals. Moreover, as the transform kernel is independent of the data, it is not efficient in terms of both decorrelation of the samples and energy compaction. Therefore, separation of the signal and noise components is generally not achievable using these suboptimal transforms. Expansion of the data into a set of orthogonal components certainly achieves maximum decorrelation of the signals. This can enable separation of the data into the signal and noise subspaces.

Principal Component Analysis (PCA) expands the data into a set of orthogonal components and certainly achieves maximum decorrelation of the signals. This can enable separation of the data into the signal and noise subspaces.

Principal components analysis on the n-by-p data matrix X returns the principal component coefficients, also known as loadings. Each principal component is a linear combination of the original variables. All the principal components are orthogonal to each other, so there is no redundant information. The principal components as a whole form an orthogonal basis for the space of the data. The PCA transformation that preserves dimensionality, that is, it gives the same number of principal components as original variables is then given by the following equation

$$Y^T = X^T W$$

$$Y^T = V \sum^T W^T W = V \sum^T$$

Principal components analysis can be defined as follows. Consider a data matrix:

$$X = \left| x_{ij} \right|$$

in which the columns represent the *p* variables and rows represent measurements of *n* objects or individuals on those variables. The data can be represented by a cloud of *n* points in a *p*-dimensional space, each axis corresponding to a measured variable. This operation defines a derived variable of the form of the following equation

$$Y_1 = a_1 x_1 + a_2 x_2 + ... + a_p x_p$$

With coefficients a_i satisfying the conditions of the following equation

$$\sum_{i=1}^{\rho} a_i^2 = 1$$

The process can be continued, until *p* mutually orthogonally is determined. Each of these lines defines a derived variable as given in the below equation

$$Y_1 = a_{1i} X_1 + a_{2i} X_2 ....a_{pi} x_p$$

where the constants a_{pi} are determined by the requirement that the variance of Y_i is a maximum, subject to the constraint of orthogonality as well as, for each *i*.

$$\sum_{k=1}^{p} a_{ik}^2 = 1$$

The Y_i thus obtained are called Principal Components of the system and the process of obtaining them is called Principal Components Analysis.

4. Sparse Representation Classifier

Sparse Representation of signals have received a vital attention in recent years [6]. Considering there are N training samples $\{ x_i , y_i \}_{i=1}^{N}$ falling into Q classes, where $x_i \in R^{d \times 1}$ is known as the i^{th} training sample and $y_i \in \{1,2,.....Q\}$ is its label.

A composite basis matrix $D = [D_{1,} D_2 ,.....D_Q]$ is desired to be found out. The requirement is that the coefficient matrix Z should be sparse column wise. If k-sparse contains the signal in which we are interested, then the sparse representation of the signal can be in any domain with non-zero coefficients. In general, SRC is always formulated and is considered as a pursuit method to optimize the objective function.

The most compact representation of a signal is being searched by the sparse representations in terms of linear combination of atoms contained in an over complete dictionary. High advancements in the multi-scale and multi-orientation representation of signals are the most important factors considered for the research on sparse representation. Sparse representation always provides a good performance when compared to that of the direct time domain processing techniques or any orthonormal transforms, and hence it is highly capable for efficient signal modeling. If the signal is corrupted, then the possible usage of the discriminative methods may fail because only little information can be obtained from discriminative analysis to deal with the noise and missing data. To address this problem, SRC always combines discrimination power, signal reconstruction and sparsity functions for the easy classification purposes. The main aim of sparse representation is to code a signal $'y'$ over a distance Φ such that $y \approx \phi\alpha$, where α is considered as a sparse vector. Always the l_0-norm is used to measure the sparse of α, which always counts the total number of non-zeroes in α. The l_0-minimization is NP-hard and therefore it cannot be widely used in sparse coding, but the l_1-minimization is the closest convex function to l_0-minimization and is widely employed in sparse coding and is represented as follows:

$$\min_{\alpha} \|\alpha\|_1 \text{ such that } \|y - \Phi\alpha\|_2 \leq \in ,$$

where $\in$ is a very small constant.

In terms of efficiency, l_1-minimization is much more efficient than l_0-minimization, but the greatest drawback associated with l_1-minimization is that, it is more time consuming

5. Results And Discussion

For sparse representation classifier based on the Performance Index, Quality values, Sensitivity, Specificity, Time and Accuracy the results are computed and tabulated in Table 5.1. The formulae for the Performance Index (PI), Sensitivity, Specificity and Accuracy are given as follows:.

$$PI = \frac{PC - MC - FA}{PC} \times 100$$

where PC – Perfect Classification, MC – Missed Classification, FA – False Alarm,

The Sensitivity, Specificity and Accuracy measures are stated by the following

$$Sensitivity = \frac{PC}{PC + FA} \times 100$$

$$Specificity = \frac{PC}{PC + MC} \times 100$$

$$Accuracy = \frac{Sensitivity + Specificity}{2} \times 100$$

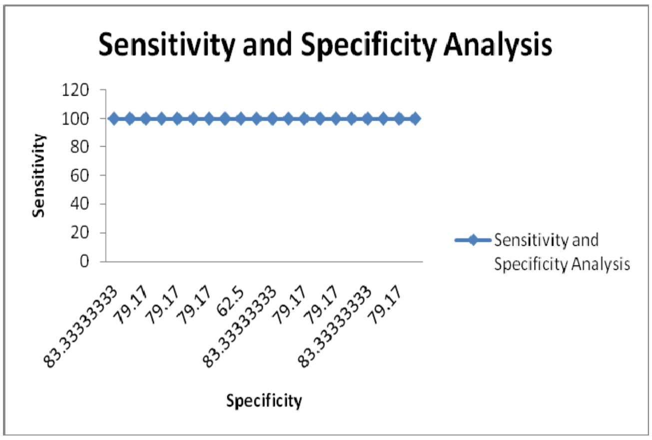

Figure 5.1 Sensitivity and Specificity Analysis when PCA acts as a dimensionality Reduction technique followed by the SRC as the Post Classifier.

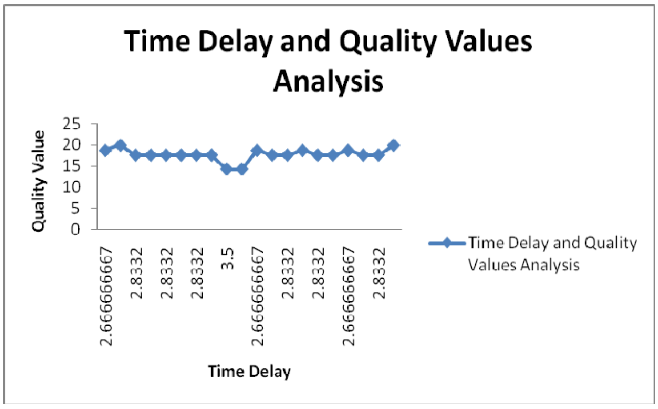

Figure 5. 2 Time Delay and Quality Value Analysis when PCA acts as a dimensionality reduction technique followed by the SRC as the Post Classifier.

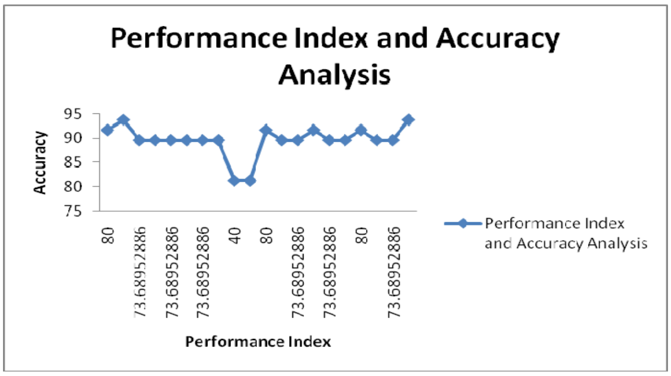

Figure 5.3 Performance Index and Accuracy Analysis using PCA as the dimensionality reduction technique and SRC as the Post Classifier

The figure 5.1 shows the sensitivity and specificity analysis using PCA as a dimensionality reduction technique and SRC as the post classifier for the perfect classification of epilepsy risk levels from EEG signals. It is inferred that the specificity remains constant throughout and there are no abrupt variations at all and the reason for this constant nature is due to the absence of the false alarm.

The figure 5.2 shows the time delay and quality value analysis when PCA acts as a dimensionality reduction technique followed by the SRC as the post classifier. The time delay is somewhat constant at certain intervals of time whereas at other time instants there is a somewhat deflection in the quality values.

From figure 5.3 it is inferred that the accuracy also shows abrupt variations throughout the performance index series and it is does not remain constant throughout.

The table 5.1 shows the average values for all the 20 patients when PCA is used as a dimensionality reduction technique and SRC is employed as a Post Classifier.

Table 5.1 Average Values for all the 20 patients when PCA is used as a dimensionality reduction technique and SRC is employed as a Post Classifier

Parameters	**Average Values for all the 20 patients**
Average Perfect Classification	79.16%
Average Performance Index	72.78%
Average Sensitivity	100%
Average Specificity	79.16%
Average Time Delay (sec)	2.833
Average Quality Value	17.76
Average Accuracy	89.58

Thus the paper gives a performance analysis by considering the Singular Value Decomposition (SVD) as a dimensionality reduction technique and Sparse Representation Classifier (SRC) as a post classifier for the perfect classification of the epilepsy risk levels obtained from Electroencephalography (EEG) signals. Performance Index (PI) and Quality Values (QV) were the two parameters that were used to assess the performance of the sparse representation classifiers. It is concluded that the average perfect classification is 79.16 % and the average performance index is 72.78%. The average sensitivity and specificity values are 100% and 79.16% respectively. The average time delay computed in seconds is obtained to be 2.833. The average Quality Value obtained in this case is 17.76 and the average accuracy obtained is 89.58 %. Future work may incorporate the usage of a variety of dimensionality reduction techniques followed by the classification of epilepsy risk levels using different types of classifiers.

REFERENCES

1. Gabor A.J (1998), ‘Seizure Detection using a Self Organizing Neural Network: Validation and Comparison with Other Detection Strategies', Electroenceph. Clin. Neuro Physiol., Vol.107, pp. 27-32.
2. Joel J. (2004), ‘Detection of Seizure Precursors from Depth EEG using a Sign Period gram Transform’, IEEE Transactions on Bio Medical Engineering, Vol.51, No.4, pp-449-458.
3. Mark van Gils, Rosen Falck A., White S., Piror P. and Gade J. (1997), ‘Signal Processing in Prolonged EEG Recordings During Intensive Care’, IEEE EMB Magazine, Vol.16, No.6, pp. 56-63.
4. Naresh C., Bhavaraju, Frei M.G. and Osorio I. (2006), ‘Analog Seizure Detection and Performance Evaluation’, IEEE Trans on BME, Vol.53, No.2, pp-238-245.
5. Srinivasan V. Eswaran C. and Sriram N. (2005), ‘Artificial Neural Network based Epileptic Detection using Time-Domain and Frequency Domain Features', Journal of Medical Systems, Vol.29, pp. 647-659
6. Bob Zhang, Vijaya kumar B.V.K, David Zhang (2014), ‘Noninvasive Diabetes Mellitus Detection Using Facial Block Color With a Sparse Representation Classifier”, IEEE Transactions on Biomedical Engineerig, Vom.61, No.4, April 2014, pp:1027-1033.