SAÜ. Fen Bil. Der. 18. Cilt, 2. Sayı, s. 93-97, 2014 SAU J. Sci. Vol 18, No 2, p. 93-97, 2014

EEG işaretlerinin epileptik analizi için boyut azaltmanın etkileri

Murat Yıldız[1*], Erhan Bergil[2], Canan Oral[3]

[1*]Sakarya Üniversitesi, Mühendislik Fakültesi, Elektrik-Elektronik Mühendisliği, Sakarya
[2]Amasya Üniversitesi, Teknik Bilimler Meslek Yüksekokulu, Elektronik ve Otomasyon Bölümü, Amasya
[3]Amasya Üniversitesi, Teknoloji Fakültesi, Elektrik-Elektronik Mühendisliği, Amasya

19.02.2013 Geliş/Received, 12.08.2013 Kabul/Accepted

ÖZET

EEG işaretleri epilepsi çalışmalarında yaygın olarak kullanılmaktadır. Bu işaretlerin özelliklerinden yararlanarak nöbet algılayan birçok yöntem önerilmiştir. Elde edilen özellik matrisi farklı sınıflandırıcılar kullanılarak sınıflandırılmaktadır. İşlem yükü özellik matrisinin boyutuyla doğrudan ilgilidir. Gerçek zamanlı uygulamalarda işlem yükünün fazla olması başlıca sorunlardandır. Bu problemi çözmek için özellik seçimi ve boyut azaltımı kullanılmaktadır. Bu çalışmada boyut azaltımının sınıflandırıcı performansları üzerindeki etkileri incelenmiştir. Sağlıklı ve epileptik bireylerden farklı koşullarda alınan EEG işaretlerinden, 300x16 boyutunda özellik matrisi elde edilmiştir. Bu matris Çok Katmanlı Yapay Sinir Ağları, Lineer Diskriminant Analizi ve Destek Vektör Makineleri yöntemleri kullanılarak sınıflandırılmıştır. Özellik matrisinin boyutu Temel Bileşenler Analiziyle 300x5 boyutuna indirgenmiştir. Sınıflandırma işlemleri boyutu indirgenmiş özellik matrisi için tekrarlanmıştır. Her iki durum için sonuçlar karşılaştırılmıştır.

Anahtar Kelimeler: EEG işaretleri, epilepsi, yapay sinir ağları, lineer diskriminant analizi, destek vektör makineleri

Effects of dimension reduction for analysis of epileptic in EEG signals

ABSTRACT

EEG signals are widely used epilepsy studies. Utilizing features of these signals, a great number of methods have been proposed for seizure detection. Obtained feature matrix is classified using different classifiers. Processing load is directly related to the size of the matrix. For real time applications, it is major problem that processing load is too much. Dimension reduction and feature selection use to eliminate this problem. In this study, effects of size reduction on classifiers' performances are investigated. A feature matrix of size 300x16 has obtained from EEG signals, which were taken from healthy and epileptic subjects in different conditions. This matrix has been classified using Multilayer Perceptron Neural Networks (MLPNN), Linear Discriminant Analysis (LDA) and Support Vector Machines (SVM). Feature matrix has been reduced to 300x5 size using Principal Component Analysis (PCA). New feature matrix has been classified using the same classifiers again. The results of both conditions have been compared.

Keywords: EEG signals, epilepsy, neural networks, linear discriminant analysis, support vector machines

* Sorumlu Yazar / Corresponding Author *myildiz@sakarya.edu.tr*

1. INTRODUCTION (GİRİŞ)

Seizures are the temporary anomalies of the brain's electrical activities. People with epilepsy who have central nervous system defect, have difficulties because of the seizures which occur at different and unforeseeable times without any symptoms. Seizures may cause convulsion and inattentiveness. In fact, seizures which occur frequently may end up with death [1].

For the diagnosis of epilepsy, many methods such as EEG, PET (Positron Emission Tomography) and MRI (Magnetic Resonant Imaging) are used. EEG signals have a common place in epilepsy studies and diagnosis [2]. EEG shows the representation of electrical activities generated by cerebral cortex neurons. For this reason, it is an important component in the clinic evaluation of brain activities, diagnosis of epilepsy and perception of epileptic attack. The signals can be obtained by placing necessary electrode to different centers and direct measures [2, 3].

A great number of features are used in order to analyze EEG signals. The mean, maximum and minimum, hjorth parameters, standard deviation and variance of EEG signals are among the features of time domain. Lyapunov exponantionals are the most widely non-linear feature. Wigner-Ville distribution may be given as an example for time-frequency distribution. Sub-frequency energy bands (Alpha, beta, theta and delta bands) are frequency domain features. Auto regressive model are among the power spectrum features [4-8].

If needed, dimension reduction can be applied by eliminating features, which haven't the distinguishing information. Principal component analysis is an example of the methods used to reduce dimension. Size reduction increases the response speed of the systems in real time applications.

After deciding on the features classification is performed. Feature matrix is given to classifier as input. At the end of classification it is aimed to separate the different components of the EEG signals into normal, pre-ictal, ictal and postictal periods.

When the literature is reviewed it is seen that Hamid R. Mohseni et al. used different features and compared the results in order to detect the seizure in EEG signals. They used separately non-linear based features, entropy based features, wavelet based features, time-frequency based features, local variance and spectral power features. Finally, they compared performances [4].

Ocak presented a new scheme for the optimal classification of epileptic seizures in EEG using wavelet analysis and the genetic algorithm (GA). Entropy values of wavelet coefficients were used as feature. The GA was used to find the optimal feature subset that maximizes the classification performance of a learning vector quantization (LVQ)-based normal and epileptic EEG classifier [9].

Shengi-Fu Liang et al. used approximate entropy, power and auto regressive model as feature. They suggested a method by which they applied genetic algorithms and principal component analysis to reduce features. In the study, linear discriminant analysis, backpropagation artificial neural networks, linear least square and support vector machine were used as classifier, the results were compared with prior studies [10].

H. Kim and J. Rosen proposed an algorithm of epileptic seizure detection for implantable device. They used auto-regressive model parameters as feature. Then they made dimension reduction using PCA [11].

Subaşı and Gürsoy classified obtained features using discrete wavelet transforms. They used principal component analysis, independent component analysis and linear discriminant analysis for dimension reduction. SVM was used as classifier. They reached 100% accuracy with linear discriminant analysis [5].

Mahajan et al. used statically features obtained from distribution of wavelet coefficients. They applied dimension reduction using Principal Component Analysis and Independent Component Analysis. Then these feature matrices were given to classifier as input. They used Artificial Neural Network as classifier and compared results [12].

2. MATERIAL AND METHOD (MATERYAL VE METOT)

The dataset used in this research are selected from the Epilepsy center in Bonn, Germany by Ralph Andrzejak [13]. We used three data sets (Set A, Set C, Set E). Each set consists of 100 pieces EEG recording, which is 23.6 second period. The first set (Set A) was taken from five healthy and awake subjects. The other sets were taken from five patients with epilepsy subjects. The second set (Set C) includes non-seizure recording. Set E includes seizure recording. All EEG signals were recorded with the same 128-channel amplifier system, using an average common reference. The data were digitized at 173.61 samples per second using 12 bit resolution.

In this study, feature matrix was obtained at 300x16 dimensions. Then, the classification was implemented using Multilayer Perceptron Neural Networks, Support Vector Machine and Linear Discriminant Analysis. Finally, the performances of classifiers were compared. 64 pieces of each set were used as training data. The others were used as testing data. Then, feature matrix was reduced to the 300x5 using Principal Component Analysis. The same procedure was repeated for reduced feature matrix. Performances were evaluated in terms of classification accuracies and elapsed times for both cases. Block scheme of process is shown figure 1.

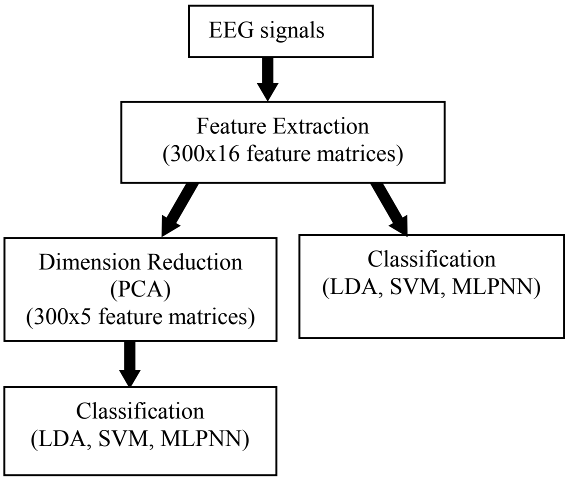

Figure 1. Block scheme of process (sürecin blok şeması)

2.1. Feature Extraction (Özellik Çıkarma)

The common features in the evaluation of EEG signals were determined and used 16 ones of these features. In this study sub-band powers (alpha, beta, delta and theta bands), total power, hjorth parameters (mobility, activity and complexity) and auto regressive coefficient (8 ones) used as feature.

2.2. Dimension Reduction (Boyut Azaltma)

Dimension reduction plays an important role in classification performance. Feature vector has usually high dimension. The aim of feature selection is to reduce dimensionality of the measurement space to a space suitable for application of classification algorithms. The feature space can be transformed space that has lower dimension than the original [14, 15].

In feature selection step, principal component analysis (PCA) is used for dimension reduction. Principal component analysis is approach to reduce dimensionality. PCA transforms a number of correlated variables in to a smaller number of uncorrelated variables called principal components [16]. First covariance matrix is calculated using feature vector. Next, eigenvectors and eigenvalues are computed and sorted according to decreasing eigenvalue. Bigger variances of data distribution are more effective to discriminate the classes than smaller variances [17].

In this study we reduced the feature matrix dimensionality from sixteen to five using Principal Component Analysis.

2.3. Classification (Sınıflandırma)

Linear Discriminant Analysis, Support Vector Machine and Multilayer Perceptron Neural Networks were used as classifier.

2.3.1. Multilayer perceptron neural networks (Çok katmanlı yapay sinir ağları)

Artificial neural networks, parallel distributed structures, because of their ability to learn and to generalize it, have the ability to solve complex problems. Neural Networks operates successfully in many engineering fields such as pattern classification, signal processing, system identification and control [18,19].

A multilayer artificial neural network consists of an input layer, one or more hidden layer and an output layer. Each layer comprises at least one neuron. Multilayer feed forward neural networks apply non-linear mapping between the input and output space. The outputs of neurons in a layer are given to next layer as input. Input layer transmits to hidden layer the information provided by the external environment without change. The information is processed in hidden and output layers, than output is defined. MLPNN passes weights assigned to different layers, and determines the output and compares it with target output. Then it propagates error signal and adjust the connection weights correspondingly [16, 18].

2.3.2. Support vector machines (Destek vektör makineleri)

SVM are a supervised learning method based on statistical learning. This method has plenty of advantages according to the traditional learning methods [20, 21]. SVM is known as a learning method based on data partitioning. The generalization errors of SVM are associated with the allocated width of the border, which is used to separate the data. SVM defines class label of unknown data via function of classification (hyper-plane) by dividing data space. For this purpose, SVM uses most suitable separator hyper-plane to divide the data [20, 22].

2.3.3. Linear discriminant analysis (Lineer diskriminant analizi)

LDA is one of the classification methods, widely used in the fields of statistics and machine learning. LDA tries to find the linear combination of features, which can be used as a linear classifier, to separate the samples belonging to two or more classes. In other words, LDA is a method; tries in order to get vectors belong to the space, which are able to separate each class [23, 24]. Discriminant analysis determines the distinctive functions. Method gets distinctive variables, which are the most dominant in distinction between groups, via these functions. Finally, it decides the group of unit [24].
This approach generates a new variable which is the combination of the available data. This variable closer to data points in the same class each other, and pushes apart the data points belonging to different classes [24].

3. EXPERIMENTAL RESULTS (DENEYSEL SONUÇLAR)

The classification process of first, feature matrix has been performed with LDA, MLPNN and SVM classifiers. SVM classifier has RBF Kernel function with σ=1,5 and regularization parameter C=100. MLPNN classifier has three layers. The input layer has sixteen neurons, hidden layer has thirty two neurons and output layer has one neuron. The results of classifiers are shown in Table 1.

Table 1. Classification results for 300x16 feature matrix. (300x16 özellik matrisi için sınıflandırma sonuçları)

The Results of LDA			
	Healthy	Non-seizure	Seizure
Healthy	36	-	-
Non-seizure	-	34	2
Seizure	2	-	34
The Results of SVM			
	Healthy	Non-seizure	Seizure
Healthy	36	-	-
Non-seizure	-	35	1
Seizure	-	-	36
The Result of MLPNN			
	Healthy	Non-seizure	Seizure
Healthy	36	-	-
Non-seizure	2	31	3
Seizure	1	-	35

The classification process of reduced feature matrix has been performed in the same way. SVM classifier has RBF Kernel function with σ=1,5 and regularization parameter C=10. MLPNN classifier has three layers. The input layer has five neurons, hidden layer has ten neurons and output layer has one neuron. The results of classifiers are shown in Table 2.

Table 2. Classification results for 300x5 feature matrix (300x5 özellik matrisi için sınıflandırma sonuçları)

The Results of LDA			
	Healthy	Non-seizure	Seizure
Healthy	32	4	-
Non-seizure	2	32	2
Seizure	2	-	34
The Results of SVM			
	Healthy	Non-seizure	Seizure
Healthy	36	-	-
Non-seizure	1	35	-
Seizure	1	-	35
The Result of MLPNN			
	Healthy	Non-seizure	Seizure
Healthy	36	-	-
Non-seizure	3	31	2
Seizure	-	1	35

SVM has the best results for both feature matrices. Dimension reduction impacts mostly on LDA's performance. The number of false detection increases from four to ten. This number hasn't changed for MLPNN. Only the periods confused with each other have changed. As a result of dimension reduction, MLPNN's performance is better than LDA's performance for healthy periods. But there is an opposite situation for non-seizure periods. LDA's performance is better than MLPNN's performance for both feature matrices. The elapsed time after dimension reduction is better than previous ones for all classifiers.

4. CONCLUSION (SONUÇLAR)

Feature matrices and classifiers used in the evaluating of EEG signals play an important role in the determining the level of success in real time applications. Dimension reduction and feature selection are often used to reduce the response time of the system. In this study, the effect of dimension reduction on the performance of the classifier has been investigated. Accuracies of classifiers are shown in Table 3. for both feature matrices.

Table 3. Accuracies of Classifiers (Sınıflayıcı Doğrulukları)

Classifiers	LDA	SVM	MLPNN
Accuracy for 300x16 feature matrix	% 96.30	% 99.07	% 94.44
Accuracy for 300x5 feature matrix	% 90.74	% 98.15	% 94.44

Classification times have been measured for each classifier. Elapsed times have been obtained considering the training and testing time. The classification times are shown in Table 4.

Table 4. Classification times (Sınıflandırma Zamanları)

Classifiers	LDA	SVM	MLPNN
Elapsed time for 300x16 feature matrix	0.1010	0.3676	47.4834
Elapsed time for 300x5 feature matrix	0.0381	0.2124	28.0576

Dimension reduction hasn't affected the accuracy of MLPNN. The elapsed times of classification, obtained using dimension reduction are better than previous ones. Designers should achieve the best performance by optimizing between classification time and accuracy.

REFERENCES (KAYNAKLAR)

[1] Shoeb, A. and Guttag, J. (2010), 'Application of Machine Learning To Epileptic Seizure Detection', The 27th International Conference on Machine Learning, Haifa, Israel, pp. 975-982.

[2] Sivasankari, N. and Thanushkodi, K. (2009) Automated Epileptic Seizure Detection in EEG Signals Using FastICA and Neural Network, Int. J. Advance, Soft Comput. Appl., Vol.1, No. 2, November, pp. 1-14.

[3] Subasi, A. (2005) Epileptic seizure detection using dynamic wavelet network, Expert Systems with Applications, vol. 29, No. 2, August, pp.343–355.

[4] Mohseni, H.R,. Maghsoudi, A. and Shamsollahi, M.B. (2006) 'Seizure Detection in EEG signals: A Comparison of Different Approaches', 28th IEEE EMBS Annual International Conference, New York City, USA, pp. 6724-6727.

[5] Subaşı, A. and Gürsoy, M.İ. (2010) EEG signal classification using PCA, ICA, LDA and support vector machines, Expert Systems with Applications, Vol. 37, No. 12, December, pp. 8659–8666.

[6] Temko, A., Thomas, E., Marnane, W., Lightbody G. and Boylan, G. (2011) EEG-based neonatal seizure detection with Support Vector Machines, Clinical Neurophysiology, Vol. 122, March, pp. 464–473.

[7] Alkan, A., Koklukaya, E. and Subasi, A. (2005) Automatic seizure detection in EEG using logistic regression and artificial neural network" Journal of Neuroscience Methods, vol. 148, No. 2, October, pp.167–176.

[8] Tzalles, A.Z., Tsipouras, M.G. and Fotiadis, D.I. (2007) Automatic Seizure Detection Based on Time-frequency analysis and Artificial Neural Networks", Computational Intelligence and Neuroscience,Vol. 2007, No. 18, August, 13 pages.

[9] Ocak, H. (2008) Optimal classification of epileptic seizures in EEG using wavelet analysis and genetic algorithm, Signal Processing, Vol. 88, No. 7, July, pp. 1858-1867.

[10] Liang, S.F., Wang, H.C. and Chang, W.L. (2010) 'Combination of EEG Complexity and Spectral Analysis for Epilepsy Diagnosis and Seizure Detection', EURASIP Journal on Advances in Signal Processing, Vol. 2010, No. 62, February.

[11] Kim, H. and Rosen, J. (2010) 'Epileptic Seizure Detection - An AR Model Based Algorithm for Implantable Device', Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE, Buenos Aires, Argentine, pp. 5541-5544.

[12] Mahajan, K.., Vargantwar, M.R. and Rajput, S.M. (2011) Classification of EEG using PCA, ICA and Neural Network, International Journal of Engineering and Advanced Technology, Vol. 1, No. 1, October, pp. 80-83.

[13] Andrzejak, R.G. (2013) EEG time series, [Online],http://www.meb.uni-onn.de/epileptologie/cms/upload/workgroup/lehnertz/eegdata.html. [1 Aug 2013]

[14] Bow, S.T. (2002) Pattern Recognition and Image Preprocessing, Marcel Dekker, New York, USA.

[15] Costaridou, L. (2005) Medical Image Analysis Methods, CRC Pres, USA.

[16] Meyer-Baese, A. (2004) Pattern Recognition for Medical Imaging, Elsevier Academic Pres, California, USA.

[17] Duda, R.O., Hart, P.E. and Stork, D.G. (2001) Pattern Classification, Wiley-Interscience, New York, USA.

[18] Haykin, S. (1999) Neural Networks: A Comprehensive Foundation, Prentice Hall, New Jersey, USA.

[19] Mitchell, T.M. (1997) Machine Learning, McGraw-Hill Science/Engineering/Math, USA.

[20] Zheng, N. and Xue, J. (2009) Statistical Learning and Pattern Analysis for Image and Video Processing, Springer-Verlag London Limited, London.

[21] Vapnik, V.N. (1998) Statistical Learning Theory, Wiley-Interscience, New York.

[22] Abe, S. (2005) Support Vector Machines for Pattern Classification, Springer, New York, USA.

[23] Junoh, A.K. and Mansor, M.N. (2012) 'Safety System Based on Linear Discriminant Analysis', 2012 International Symposium on Instrumentation & Measurement, Sensor Network and Automation (IMSNA), Vol:1, pp. 32-34.

[24] Fielding, A. (2000) Cluster and Classification Techniques for the Biosciences, Cambridge University Press.