Original Article / Özgün Araştırma

The Comparison of Different Dimension Reduction and Classification Methods in Electroencephalogram Signals

Elektroensefalografi Sinyallerinde Farklı Boyut İndirgeme ve Sınıflandırma Yöntemlerinin Karşılaştırılması

Hakan Öztürk[1], Mevlüt Türe[1], Nefati Kıylıoğlu[2], İmran Kurt Ömürlü[1]

[1]Aydın Adnan Menderes University Faculty of Medicine, Department of Biostatistics, Aydın, Turkey
[2]Aydın Adnan Menderes University Faculty of Medicine, Department of Clinical Neurology, Aydın, Turkey

Keywords
Electroencephalogram, discrete wavelet transformation, Principal component analysis, Independent component analysis, Support vector machine, Linear discriminant analysis

Anahtar Kelimeler
Elektroensefelogram, ayrık dalgacık dönüşümü, temel bileşenler analizi, bağımsız bileşen analizi, destek vektör makinesi, doğrusal ayırma analizi

Received/Geliş Tarihi : 14.06.2017
Accepted/Kabul Tarihi : 19.10.2017

doi:10.4274/meandros.96168

Address for Correspondence/Yazışma Adresi:
Mevlüt Türe MD,
Aydın Adnan Menderes University Faculty of Medicine, Department of Biostatistics, Aydın, Turkey
Phone : +90 256 444 12 56
E-mail : mevlutture@gmail.com

ORCID ID: orcid.org/ 0000-0003-3187-2322

Abstract

Objective: Electroencephalogram (EEG) signals have been broadly utilized for the diagnosis of epilepsy. Expert physicians must monitor long-term EEG signals that is sometimes difficult and time consuming process for epilepsy diagnosis. In this study, classification performances of support vector machine (SVM) and linear discriminant analysis (LDA), which are widely used in computer supported epilepsy diagnosis, were compared by using wavelet-based features of extracted from EEG signals which were derived in either normal or inter-ictal periods. In addition, principal component analysis (PCA) and independent component analysis (ICA) were used to determine the effects of dimension reduction on classification success.

Materials and Methods: The EEG data were sampled from the EEG laboratory of the Department of Neurology and Clinical Neurophysiology in Adnan Menderes University. Study was approved by Local Ethics Committee with protocol number 2016/873. Ten patients with epilepsy and 10 normal were the study group. EEG signals of patients with epilepsy were contains only seizure free- epochs. EEG signals were first decomposed into frequency sub-bands by using discrete wavelet transform (DWT) and then some statistical features were calculated from those to classify it's as normal or epileptic.

Results: In classification of the EEG signals, it's as normal or epileptic, we achieved 88.9% accuracy rate using SVM with radial basis function (RBF) kernel without dimension reduction.

Conclusion: Results showed that SVM was a powerful tool in classifying EEG signals if it's normal or epileptic.

Öz

Amaç: Bu çalışmada, epileptik ve epileptik olmayan elektroensefalografi (EEG) sinyallerinden elde edilen öznitelikler boyutlarının temel bileşenler analizi ve bağımsız bileşenler analizi yöntemleri ile indirgenmesinin sınıflandırma başarısı üzerine etkilerinin belirlenmesi ve doğrusal ayırma analizi ile destek vektör makinesi (DVM) yöntemlerinin sınıflandırma performanslarının karşılaştırılması amaçlanmıştır.

Gereç ve Yöntemler: Çalışmaya 10 kontrol ve uzman hekim tarafından epilepsi tanısı konmuş 10 hasta olmak üzere toplam 20 kişi dahil edildi. Epilepsi tanısı konmuş

hastalardan alınan EEG kayıtları nöbet geçirmedikleri sırada alınan kayıtlardı. Epileptik ve epileptik olmayan sinyalleri sınıflandırmak için ayrık dalgacık dönüşümü ile sinyallerinin spektral analizi gerçekleştirildi ve sınıflandırmada kullanılacak olan öznitelikler elde edildi. Öncelikle özniteliklerin boyutu indirgenmeden, daha sonra temel bileşenler analizi ve bağımsız bileşenler analizi ile indirgenerek sınıflandırma yapıldı. Sınıflandırma doğrusal diskriminant analizi, lineer ve radyal tabanlı çekirdek fonksiyonlarının kullanıldığı DVM yöntemleri ile gerçekleştirildi.
Bulgular: EEG sinyallerinin epileptik ya da normal olarak sınıflandırılmasında radyal tabanlı çekirdek fonksiyonunun kullanıldığı DVM ile %88,9 doğruluk oranı elde edildi.
Sonuç: DVM yönteminin, epileptik ve normal sinyalleri ayırt etmede kullanılabilecek güçlü bir yöntem olduğu sonucuna varıldı.

Introduction

Electroencephalogram (EEG) is a recording of brain electrical oscillations from the scalp by using surface electrodes (1). Clinically, EEG is particularly helpful for epilepsy. In addition to clinical history and imaging studies, EEG findings help us to determine the diagnosis and the types of epilepsy (1,2). It is also helpful for the treatment of the disease (3). Epilepsy goes with seizures that occur in a sudden and unexpected nature. Mostly, seizures don''t observe during the event by the medical staff. When they are not observed or having described an atypical features, EEG helps us to reveal whether it is a seizure or not. In general, analysis of EEG signals done by expert physicians by visual analysis. Visual recognition of epileptic waveforms is sometimes difficult and time consuming for physicians who especially have not got enough expertise (4). There is also inter-reader differencesare also inter-reader differences during the visual analysis and it suggests that the visual analysis could be insufficient. With that reason, new computer evaluation techniques are developed and performed in healthy and diseased individuals (5-13). Most of these studies consist of two steps: feature extraction from the EEG signals and then classification of these features. In many study, extracted signals are derived at the time of seizure and at normal periods. So, classification performance of these studies were performance of these studies was quite high (9,11-21). In fact, visual assessment might be sufficient at that time. In that point, the question arise that whether there is an advanced statistical techniques clearly differentiate the normal and the patients with epilepsy while there is no seizure. In this study, our aim was to classify normal and epileptic patients by using their EEG data sets that derived from the archives of patients who known as epileptic (without a seizure activity at that time) and normal.

Materials and Methods

The EEG data were sampled from the EEG laboratory of the Department of Neurology and Clinical Neurophysiology in Adnan Menderes University. Study was approved by Adnan Menderes University Local Ethics Committee with (protocol approval number: 2016/873). The EEG data were recorded by Micromed EEG device (16 channels). Ten patients with epilepsy (5 male - 5 female, mean age 34±4 years) and 10 normal (5 male - 5 female, mean age 35±5 years) were the study group. EEG signals of patients with epilepsy were contains only seizure free- epochs. Nine mm, round, golden-cup electrodes were placed according to 10-20 international electrode placement system (Figure 1). Sampling frequency was 256 Hz. Reference montage (A1 and A2) measurements were derived from each channels and the duration of epoch was 30 seconds. In both group, each EEG datum was added one another and a single EEG data that was 300 seconds long were obtained. Thus each

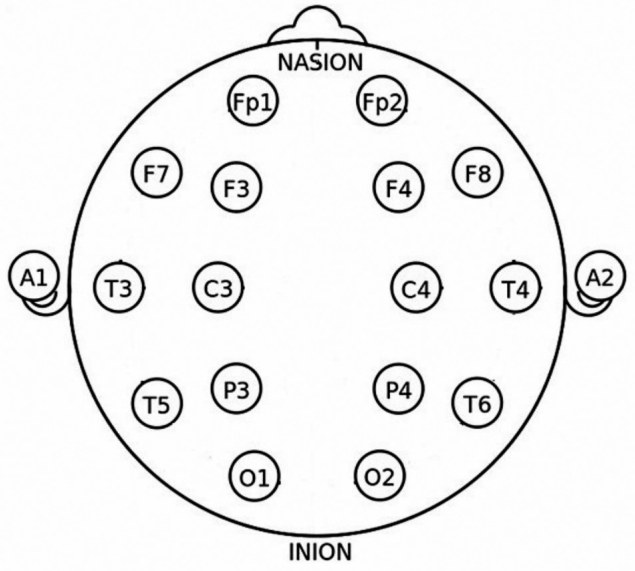

Figure 1. Electrode placement for 16 channels

channel consisted of total 76800 samples and then for each of those channels, 30 rectangular windows were formed which consists of 256 discrete data. Finally; total of 600 EEG segments, 300 epileptic and 300 normal, were obtained. EEG data was retrospectively collected.

2.1 Extracting Features with Discrete Wavelet Transform

Most of the biological signals like EEG are non-stationary signals. In other words, the amplitude, phase, and frequency of EEG signals are constantly changing signals. Various methods are used to analyse changes in the EEG signal (5). Wavelet transform (WT) is one of the most common methods which is used for time-frequency analysis of EEG signals. WT provides optimum time-frequency resolution over all frequency ranges (22). Therefore it has been widely used to provide a quantitative measure of the frequency distribution of the EEG and detect the presence of particular patterns (6).

WT analysis can be classified as two types: Continuous wavelet transform (CWT) and discrete WT (DWT). CWT is obtained by taking a projection of the signal to the functions created by scaling and shifting of a mother wavelet function. The mother wavelet function is a prototype function used to generate wavelets (23). According to the principle, the CWT is defined as:

$$CWT(s,\tau) = \frac{1}{\sqrt{s}} \int_{-\infty}^{+\infty} x(t)\psi\left(\frac{t-\tau}{s}\right) dt$$

where x(t), ψ, s and τ denote the signal to be processed, the wavelet function, scaling and shifting parameters, respectively. The different window functions used for the transformation are derived by shifting and scaling the mother wavelet. The shifting parameter τ changes the position of the window function on the signal. Hence the window moves on the signal. The scale parameter s expands or contracts the window function. Large values of s are suitable for general views and small values are suitable for detailed views. 1/√s is the normalization multiplier that ensures that the energy is the same for all values of s (13,22).

Calculation of the wavelet coefficients for every possible scale causes unnecessary information to be received from the signal. Moreover it takes a long time (24). If the scaling and shifting parameters are chosen as powers of 2, the analysis becomes more effective and faster (25). This method is called DWT and can be defined as:

$$DWT(j,k) = \frac{1}{\sqrt{2^j}} \int_{-\infty}^{+\infty} x(t)\psi\left(\frac{t-2^j k}{2^j}\right) dt$$

where the parameters s and τ are replaced by 2^j and 2^j k.

In the DWT, the signal is decomposed into approximation and detail coefficients at the first level by using low and high pass filters (Figure 2). Then the approximation coefficients are further decomposed into next level of approximation and detail coefficients (26,27).

It is very important to determine the appropriate wavelet function and the level of decomposition. The level of decomposition is chosen based on the dominant frequency components of the signal (13).

In this study DWT was employed to decompose the EEG signals into different frequency bands for different wavelet functions. Due to its high success the Daubechies wavelet order in 4 (Db4) was used to construct the feature vectors (13,15). Since the EEG signals do not have any useful frequency components above 30 Hz, the number of levels was chosen to be 6. After decomposition, D1-D6 details and A6 approximation coefficients were obtained (Figure 3).

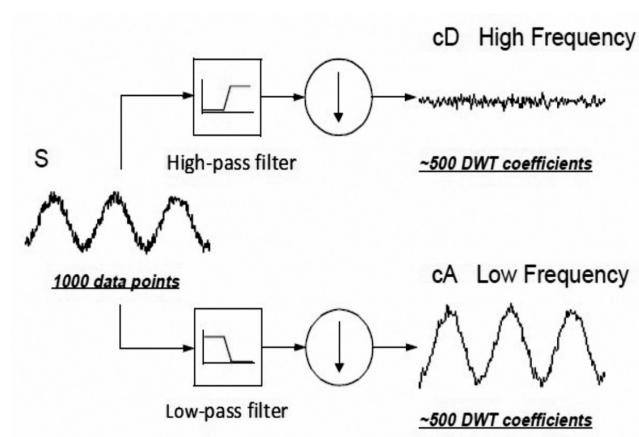

Figure 2. Decomposition of a signal into approximation and detail coefficients
DWT: Discrete wavelet transform

The extracted wavelet coefficients give a compact illustration that shows the energy distribution of the EEG signal in time and frequency. To characterize the EEG signals, statistics over the set of the wavelet coefficients were calculated (5, 6, 13, 27, 28). The following statistics were calculated from the wavelet coefficients:

1. Mean of the absolute values of the coefficients in each sub-band $\left(\frac{\sum|x_i|}{n}\right)$

2. The minimum value of the coefficients in each sub-band $\left(\min(x_i)\right)$.

3. The maximum value of the coefficients in each sub-band $\left(\max(x_i)\right)$.

4. Standard deviation of the coefficients in each sub-band $\left(s = \sqrt{\frac{\sum(x_i-\bar{x})^2}{n-1}}\right)$

5. Entropy of the coefficients in each sub-band ($S=\sum p_x \log p_x$).

6. Energy of the coefficients in each sub-band ($E=\sum|x_i|^2$).

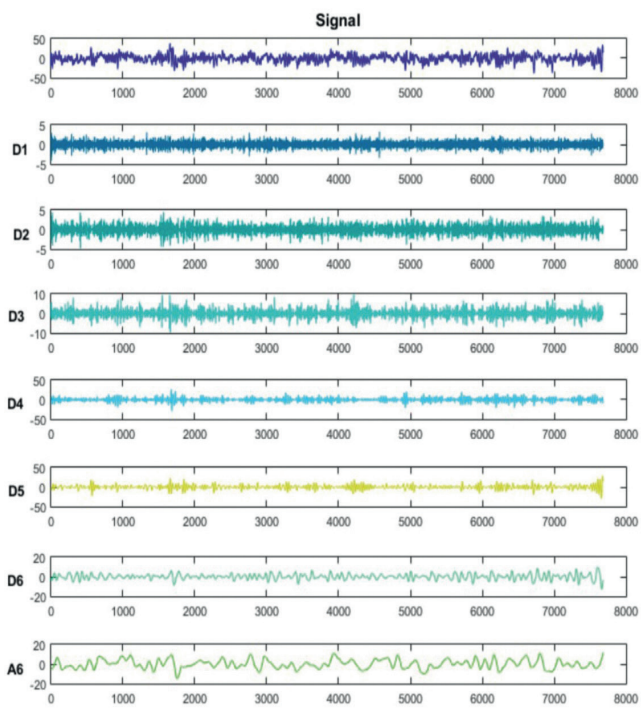

Figure 3. Approximate and detailed coefficients of epileptic electroencephalogram signal after decomposition

Since the frequency components above 30 Hz is lack of use in epilepsy analysis, the features were extracted from D4 (16-32 Hz), D5 (8-16 Hz), D6 (4-8 Hz) detail coefficients and A6 (0-4 Hz) approximation coefficients. Thus 24 statistical features were obtained from each channel. In total 384 features were obtained from 16 channels and normalized in [0,1].

Dimensionality Reduction of Features

Principal Component Analysis

Principal component analysis (PCA) is a transformation technique that reduces the dimension of p-dimensional data set containing correlated variables to a lower dimensional space containing uncorrelated variables while preserving the existing variability in the data set as much as possible. The variables obtained by the transformation are called the principal components of the original variables. The first principal component captures the maximum variance in the data set and the others capture the remaining variance according to decreasing order (29,30).

The number of principal components that can be obtained for p number of variables is at most p, and the principal components are formed as linear combinations of variables (29). A linear combination of any **x** random vector can be expressed as:

$$\mathbf{y_1} = \sum^{p} a_{k1}\, x_k = \mathbf{a}_1'\mathbf{x}$$

where $a_{11}, a_{21},\ldots,a_{p1}$ are the weighting coefficients of the weight vector $\mathbf{a_1}$ and $\mathbf{y_1}$ represents the first principal component. The variance of $\mathbf{y_1}$ depends on the norm and direction of $\mathbf{a_1}$. As the norm of a_1 increases, the variance of $\mathbf{y_1}$ will also increases. Therefore it is aimed to obtain the maximum variance by introducing a constraint such that the norm of $\mathbf{a_1}$ is 1. Under this constraint, the variance of the first principal component expressed as (29):

$$Var(\mathbf{a}_1) = E[\mathbf{y}_1^2] = \mathbf{a}_1' E[\mathbf{x}\mathbf{x}']\mathbf{a}_1 = \mathbf{a}_1'\mathbf{C}_x\mathbf{a}_1$$

where $\mathbf{C_x}$=E[**xx'**] denotes covariance matrix. The result that maximizes Var($\mathbf{a_1}$) is obtained by calculation of eigenvectors $\mathbf{v_1},\ldots,\mathbf{v_n}$ corresponding to eigenvalues $\lambda_1,\ldots,\lambda_n$ ($\lambda_1 \geq \cdots \geq \lambda_n$) of $\mathbf{C_x}$ matrix. The first principal component is expressed as (29):

$$\mathbf{y}_1 = \mathbf{v}_1'\mathbf{x}$$

The second principal component with the constraint that is uncorrelated to the first principal component ($E[\mathbf{y}_1\mathbf{y}_2]=0$) is expressed as (29):

$$\mathbf{y}_2 = \mathbf{v}_2'\mathbf{x}$$

By this way, the mth principal component such that $1 \leq m \leq p$ and $E[\mathbf{y}_k\,\mathbf{y}_m]=0$ ($k \neq m$) is expressed as (29):

$$\mathbf{y}_m = \mathbf{v}_m'\mathbf{x}$$

PCA has been frequently used in studies on epilepsy diagnosis with EEG signals (13, 31, 32).

Independent Component Analysis

Independent component analysis (ICA) is a statistical method that tries to distinguish between multiple randomly mixed signals without knowing the mixing mechanism. ICA assumes that each measured signal is a linear combination of independent signals. It decomposes multidimensional data vector linearly to statistically independent components (5). The mixing model can be written as:

$$x=As$$

where x denotes the random vector whose elements are the mixtures $x_1,\ldots,x_n$, and s denotes the vector of the original source signals with elements $s_1,\ldots,s_n$, and A denotes the mixing matrix with elements a_{ij} (33). In equation (8) neither A nor s are known. If a matrix W can be found as the inverse of of A, the original source signals can be estimated. The estimated signals can be expressed as:

$$y=Wx$$

where y denotes the vector whose elements are the estimations of the original source signals (33). A number of algorithms have been developed for estimating W. One of these algorithms is the fast fixed-point algorithm (FastICA) developed by Hyvärinen (34). FastICA provides fast convergence, easy to apply and reliable results. In this study FastICA algorithm was used to estimate the W.

Classification of Features

Linear Discriminant Analysis

The goal of linear discriminant analysis (LDA) is to derive a discriminant function to maximize the difference between the groups. In LDA, the number of discriminant functions is determined according to the number of the groups. If there are two groups, then one discriminant function is used. A discriminant function consists of a linear combination of predictors. The weights of the predictors are calculated such that the ratio of the variance between classes to the variance within class is maximized. The discriminant function for two groups and p predictors expressed as:

$$D=w_0+w_1 X_1+w_2 X_2+\cdots\cdots+w_p X_p$$

where w_0, w_i and X_i ($i = 1,\ldots,p$) denote the constant, the weights of the predictors and the predictors, respectively.

Support Vector Machine

In machine learning, support vector machines (SVMs) are supervised learning models with associated learning algorithms that analyse data used for classification and regression analysis (35).

SVM aims to find the best separating hyper plane (optimal hyper plane), with the maximum distance between observations in the two classes. The basic support vector classifier for linear separable data is shown in Figure 4. Where w is the normal of optimal hyper plane, b is bias and x is the features vector. The optimal hyper plane ($\mathbf{w}^T\mathbf{x}+b=0$) divides the plane into two sets depending on the sign of $\mathbf{w}^T\mathbf{x}+b$ (36).

SVM maps the data that cannot be separated linearly into a higher dimensional space in which they can be separated linearly by using an appropriate kernel function (Figure 5) (9, 35, 37).

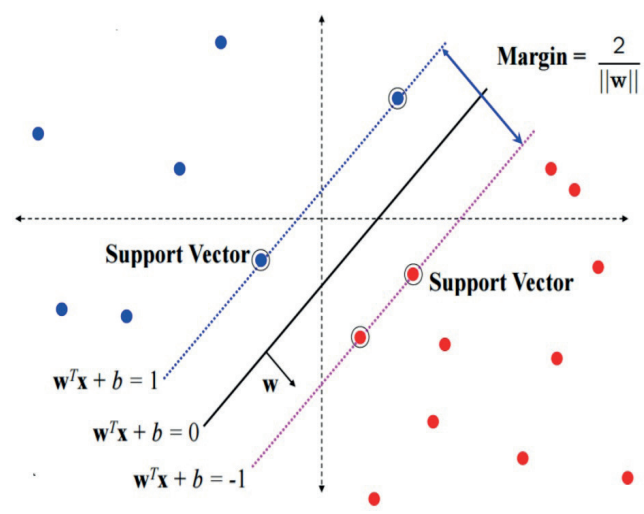

Figure 4. The linear support vector classifier

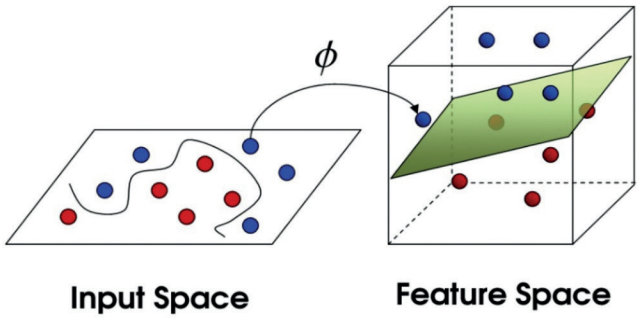

Figure 5. Non-linear support vector classifier

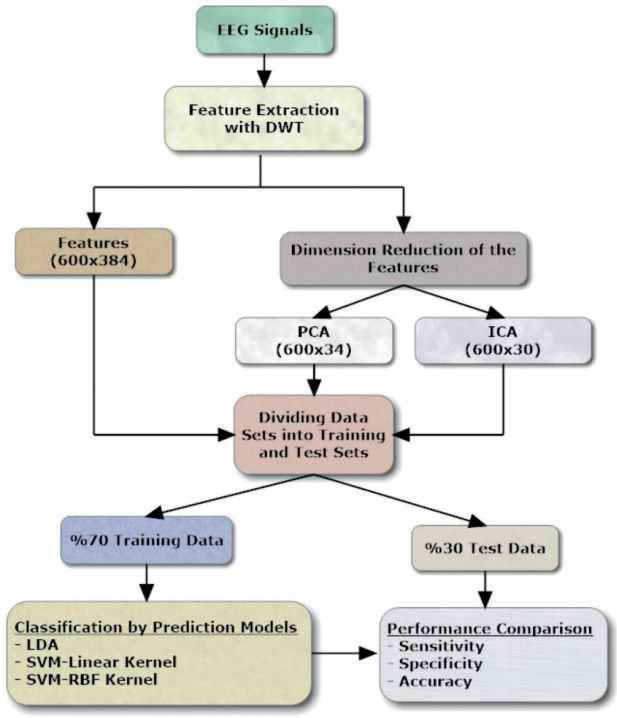

Figure 6. Diagram of the application process
EEG: Electroencephalogram, DWT: Discrete wavelet transform, SVM: Support vector machine, LDA: Linear discriminant analysis, PCA: Principal component analysis, ICA: Independent component analysis

In this study, the linear kernel: $K(x_i,x_j)=x_i.x_j$ and the radial basis function (RBF) kernel: $K(x_i,x_j)=\exp(-\gamma\ ||x_i-x_j||^2)$ were used as the kernel functions.

The hyper parameters C and γ of the SVM classifier were determined through a 10-fold cross-validation grid search performed and the parameters giving the highest accuracy were used in SVM classifiers.

Results

A comparison of both the effects of PCA and ICA on the classification performance and the classification performances of LDA and SVM with linear and RBF kernels was done on the EEG data that derived from the normal and patients with epilepsy. Classifications were performed for three different feature matrices: (a) features without dimension reduction (384 features), (b) features being reduced by PCA (34 features) and the and (3) features being reduced by ICA (30 features). The eigenvalues-greater-than-one rule proposed by Kaiser (1960) was used to determine for the number of components. These three feature matrices were classified by using LDA, SVM with linear kernel and RBF kernel. Sensitivity, specificity and accuracy rates were used for the performance measures of the classifiers. Additionally, for all feature matrices used in the classifications, training and test data sets were randomly divided into two parts: 70% training (n=420) and 30% test (n=180) data corresponding to the same points. The process performed in the application is given in Figure 6.

In training sets; among the classifiers, SVM with RBF kernel reached to the highest accuracy rate (96.2%) at the features without dimension reduction, SVM with linear kernel reached to the highest at the (94.7%) at similar to the previous; LDA reached to the highest rate (79.8%) at the features reduced by PCA (Table 1).

In test sets, the highest accuracy (88.9%) was obtained by SVM with RBF kernel without dimension reduction. The highest accuracy (82.2%) for SVM with linear kernel was obtained with the features being reduced by PCA. The highest accuracy (78.9%) for LDA was obtained for the features being reduced by both PCA and ICA (Table 2).

Discussion

A number of studies have been done to classify EEG signals for the diagnosis of epilepsy. A direct comparison of the previous studies that using EEG signals is hard due to the variety of EEG datasets, wavelet types, decomposition levels and also the variety of the statistical features used in the classification process (15). Previously, many researchers used the same EEG data which included five sets (named as A-E) described by Andrzejak, et al. (38). Set A and B were obtained from normal, set C-E were obtained from patients with epilepsy. Set C and D were included seizure-free interval while set E were include seizure-related interval. Wavelet-based

Table 1. The classification performances of prediction models in training sets

Training data sets	Classifiers	Sensitivity (%)	Specificity (%)	Accuracy (%)
Features without dimension reduction	LDA	69.0	83.3	76.2
	SVM-Linear kernel	91.4	98.1	94.7
	SVM-RBF kernel	93.3	99.1	96.2
Features reduced by PCA	LDA	73.8	85.7	79.8
	SVM-Linear kernel	80.0	88.6	84.3
	SVM-RBF kernel	83.8	90.9	87.4
Features reduced by ICA	LDA	70.5	86.2	78.3
	SVM-Linear kernel	80.0	85.2	82.6
	SVM-RBF kernel	74.8	93.8	84.3

SVM: Support vector machine, LDA: Linear discriminant analysis, PCA: Principal component analysis, ICA: Independent component analysis

Table 2. The classification performances of prediction models in test sets

Test data sets	Classifiers	Sensitivity (%)	Specificity (%)	Accuracy (%)
Features without dimension reduction	LDA	57.8	63.3	60.6
	SVM-Linear kernel	77.7	81.1	79.4
	SVM-RBF kernel	92.2	85.6	88.9
Features reduced by PCA	LDA	67.8	90.0	78.9
	SVM-Linear kernel	78.9	85.6	82.2
	SVM-RBF kernel	83.3	82.2	82.8
Features reduced by ICA	LDA	71.1	86.6	78.9
	SVM-Linear kernel	78.8	84.4	81.7
	SVM-RBF kernel	80.0	88.9	84.4

SVM: Support vector machine, LDA: Linear discriminant analysis, PCA: Principal component analysis, ICA: Independent component analysis

features that obtained from these sets (Set E and the others) were used to assess the performances of classifiers and to detection of a seizure activity. Nearly or exactly 100% accuracy rates were obtained with different classification methods. Among them, Xie and Krishnan (39) used k-nearest neighborhood method, Kumar et al. (11) used artificial neural network, Das et al. (9) used SVM method and all of these studies accuracy rates were 100%. In these sets, there were significant differences between the signals that were derived at the time of seizure and at normal periods, and no need to use of complicated statistical methods instead of using a simple threshold value or even visual assessment. It was also difficult to say that any classification method was superior to other methods. Different from these studies, we did not use an epileptic seizure activity in our EEG data set in patients with epilepsy. As a result, classification attempts of our data were more difficult than the previous data set and showed the real discrimination ability of these methods.

Orhan et al. (40) classified the wavelet-based features that obtained from data set A and D described by Andrzejak, et al. (38) and obtained 96% accuracy rate with a multilayer perception neural network model. Subasi and Ercelebi (41) used similar data sets that we used in this study. They classified the wavelet-based features and obtained respectively 93.0% and 89.0% accuracy rates for artificial neural network and logistic regression. In this study, we also classified some wavelet-based features and got respectively 88.9%, 82.2% and 78.9% accuracy rates for SVM with RBF kernel, SVM with linear kernel, and LDA. One of the most important reasons for different accuracy rates among the studies was the using of different data sets. The ratio of epileptic abnormalities in EEG data sets was variable and it made difficult to compare of the classification studies.

There is still a debate that linear or non-linear method has more successful than each others for classifying of EEG signals. Garrett et al. (42) classified five different mental states using EEG signals and achieved respectively 66.0%, 69.4% and 72.0% accuracy rates for LDA, artificial neural network and SVM with RBF kernel. They favoured that non-linear methods were more successful than linear methods for the classification of EEG signals. Lehmann, et al. (43) did the same classification of the EEG data that derived from the Alzheimer patients and normal, and they got 91.0% and 95.0% accuracy rates with LDA and SVM with RBF kernel. Supriya et al. (19) classified the EEG data sets A and E [described by Andrzejak et al. (38)] and achieved 86.87%, 99.25% and 100% accuracy rates with LDA, SVM with linear kernel and SVM with polynomial kernel. In this study, we achieved respectively 78.9%, 82.2% and 88.9% accuracy rates with LDA, SVM with linear kernel and SVM with RBF kernel. In this context, we have achieved similar results with previous studies.

Subasi and Gursoy (13) used PCA, ICA and LDA to reduce the dimension of the features that obtained from EEG signals and compared the performances on classification success. They used SVM with RBF kernel to classify the reduced features and achieved respectively 98.7%, 99.5% and 100% accuracy rates for PCA, ICA and LDA. We used SVM with RBF kernel to classify the reduced features and for ICA we found the highest accuracy rate 84.4% and for PCA the highest accuracy rate 82.8%. Hence we found similar result with Subasi and Gursoy (13). In addition, with dimensionally reduction, both the performance of LDA and SVM with linear kernel were increased. The highest classification performances in all data sets were got by using SVM with RBF kernel. It achieved the highest accuracy rate (88.9%) for the features without dimension reduction. It was noted that the dimension reduction methods were adversely affected the performance of its.

Conclusion

In this study, the classification performances of SVM and LDA, which are widely used for computer supported diagnose of epilepsy, were compared by using wavelet-based features extracted from EEG signals. In addition, PCA and ICA were used to determine the effects of dimension reduction on the classification success. Results showed that, SVM with RBF kernel achieved the highest accuracy rate (88.9%) for the features without dimension reduction. The dimension reduction methods PCA and ICA improved classification performances of LDA and SVM with linear kernel, but decreased the classification performance of SVM with RBF kernel. Consequently, with dimensionally reduction, LDA and SVM with linear kernel perform better classifications.

Ethics

Ethics Committee Approval: Adnan Menderes University Ethics Committee. (approval no: 2016/873).

Informed Consent: It was not taken.

Peer-review: Externally peer-reviewed.

Authorship Contributions

Surgical and Medical Practices: N.K., Concept: M.T., H.Ö., Design: M.T., İ.K.Ö., H.Ö., Data Collection or Processing: N.K., M.T., H.Ö., Analysis or Interpretation: M.T., H.Ö., İ.K.Ö., Literature Search: M.T., H.Ö. Writing: H.Ö.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

References

1. Nunez PL, Srinivasan R. Electric fields of the brain: the neurophysics of EEG: Oxford University Press, USA; 2006.
2. Ropper AH. Adams and Victor's principles of neurology: McGraw-Hill Medical Pub. Division New York; 2005.
3. Sanei S, Chambers J. EEG Signal processing, 2007. John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England; 2007.
4. Nuh M, Jazidie A, Muslim M, editors. Automatic detection of epileptic spikes based on wavelet neural network. Circuits and Systems, 2002 APCCAS'02 2002 Asia-Pacific Conference on; 2002: IEEE.
5. Acharya UR, Sree SV, Swapna G, Martis RJ, Suri JS. Automated EEG analysis of epilepsy: a review. Knowledge-Based Systems 2013; 45: 147-65.
6. Amorim P, Moraes T, Fazanaro D, Silva J, Pedrini H. Electroencephalogram signal classification based on shearlet and contourlet transforms. Expert Syst Appl 2017; 67: 140-7.
7. Bao FS, Lie DY-C, Zhang Y, editors. A new approach to automated epileptic diagnosis using EEG and probabilistic neural network. Tools with Artificial Intelligence, 2008 ICTAI'08 20th IEEE International Conference on; 2008: IEEE.
8. Bronzino JD. Biomedical engineering handbook: CRC press; 1999.
9. Das AB, Bhuiyan MIH, Alam SS. Classification of EEG signals using normal inverse Gaussian parameters in the dual-tree complex

wavelet transform domain for seizure detection. Signal Image and Video Processing 2016; 10: 259-66.
10. Fu K, Qu J, Chai Y, Zou T. Hilbert marginal spectrum analysis for automatic seizure detection in EEG signals. Biomed Signal Process Control 2015; 18: 179-85.
11. Kumar Y, Dewal M, Anand R. Epileptic seizures detection in EEG using DWT-based ApEn and artificial neural network. Signal Image and Video Processing 2014; 8: 1323-34.
12. Sharma R, Pachori RB. Classification of epileptic seizures in EEG signals based on phase space representation of intrinsic mode functions. Expert Systems with Applications. 2015; 42: 1106-17.
13. Subasi A, Gursoy MI. EEG signal classification using PCA, ICA, LDA and support vector machines. Expert Syst Appl 2010; 37: 8659-66.
14. Alam SM, Bhuiyan MI. Detection of seizure and epilepsy using higher order statistics in the EMD domain. IEEE J Biomed Health Inform 2013; 17: 312-8.
15. Amin HU, Malik AS, Ahmad RF, Badruddin N, Kamel N, Hussain M, et al. Feature extraction and classification for EEG signals using wavelet transform and machine learning techniques. Australas Phys Eng Sci Med 2015; 38: 139-49.
16. Iscan Z, Dokur Z, Demiralp T. Classification of electroencephalogram signals with combined time and frequency features. Expert Systems with Applications 2011; 38: 10499-505.
17. Lima CA, Coelho AL, Eisencraft M. Tackling EEG signal classification with least squares support vector machines: a sensitivity analysis study. Comput Biol Med 2010; 40: 705-14.
18. Srinivasan V, Eswaran C, Sriraam N. Artificial neural network based epileptic detection using time-domain and frequency-domain features. J Med Syst 2005; 29: 647-60.
19. Supriya S, Siuly S, Zhang Y. Automatic epilepsy detection from EEG introducing a new edge weight method in the complex network. Electronics Letters 2016; 52: 1430-2.
20. Tzallas AT, Tsipouras MG, Fotiadis DI. Epileptic seizure detection in EEGs using time–frequency analysis. IEEE Trans Inf Technol Biomed 2009; 13: 703-10.
21. Zhang T, Chen W. LMD based features for the automatic seizure detection of EEG signals using SVM. IEEE Trans Neural Syst Rehabil Eng 2017; 25: 1100-8.
22. Steinbuch M, van de Molengraft M. Wavelet theory and applications: a literature study. Eindhoven: Eindhoven University Technology Department of Mechanical Engineering Control System Group. 2005.
23. Kovacevic MVJ. Wavelets and subband coding. 1995.
24. Rioul O, Vetterli M. Wavelets and signal processing. IEEE signal processing magazine 1991; 8: 14-38.
25. Misiti M, Misiti Y, Oppenheim G, Poggi J-M. Matlab Wavelet Toolbox User\'s Guide. Version 3. 2004.
26. Faust O, Acharya UR, Adeli H, Adeli A. Wavelet-based EEG processing for computer-aided seizure detection and epilepsy diagnosis. Seizure 2015; 26: 56-64.
27. Li M, Chen W, Zhang T. Classification of epilepsy EEG signals using DWT-based envelope analysis and neural network ensemble. Biomed Signal Process Control 2017; 31: 357-65.
28. Meier R, Dittrich H, Schulze-Bonhage A, Aertsen A. Detecting epileptic seizures in long-term human EEG: a new approach to automatic online and real-time detection and classification of polymorphic seizure patterns. J Clin Neurophysiol 2008; 25: 119-31.
29. Dunteman GH. Principal component analysis (Quantitative applications in the social sciences). SAGE Publications, Thousand Oaks; 1989.
30. Jolliffe I. Principal component analysis: Wiley Online Library. 2005.
31. Acharya UR, Sree SV, Alvin APC, Suri JS. Use of principal component analysis for automatic classification of epileptic EEG activities in wavelet framework. Expert Syst Appl 2012; 39: 9072-8.
32. Ghosh-Dastidar S, Adeli H, Dadmehr N. Principal component analysis-enhanced cosine radial basis function neural network for robust epilepsy and seizure detection. IEEE Trans Biomed Eng 2008; 55: 512-8.
33. Hyvärinen A, Karhunen J, Oja E. Independent Component Analysis, 500 John Wiley & Sons. NJ, USA. 2001;501.
34. Hyvärinen A. Fast and robust fixed-point algorithms for independent component analysis. IEEE Trans Neural Netw 1999; 10: 626-34.
35. Cortes C, Vapnik V. Support-vector networks. Machine learning. 1995; 20: 273-97.
36. Ben-Hur A, Weston J. A user's guide to support vector machines. Methods Mol Biol 2010; 609: 223-39.
37. Burges CJ. A tutorial on support vector machines for pattern recognition. Data Min Knowl Discov 1998; 2: 121-67.
38. Andrzejak RG, Lehnertz K, Mormann F, Rieke C, David P, Elger CE. Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state. Physical Review E 2001; 64: 061907.
39. Xie S, Krishnan S. Wavelet-based sparse functional linear model with applications to EEGs seizure detection and epilepsy diagnosis. Med Biol Eng Comput 2013; 51: 49-60.
40. Orhan U, Hekim M, Ozer M. EEG signals classification using the K-means clustering and a multilayer perceptron neural network model. Expert Syst Appl 2011; 38: 13475-81.
41. Subasi A, Ercelebi E. Classification of EEG signals using neural network and logistic regression. Comput Methods Programs Biomed 2005; 78: 87-99.
42. Garrett D, Peterson DA, Anderson CW, Thaut MH. Comparison of linear, nonlinear, and feature selection methods for EEG signal classification. EEE Trans Neural Syst Rehabil Eng 2003; 11: 141-4.
43. Lehmann C, Koenig T, Jelic V, Prichep L, John RE, Wahlund L-O, et al. Application and comparison of classification algorithms for recognition of Alzheimer's disease in electrical brain activity (EEG). J Neurosci Methods 2007; 161: 342-50.