Engineering Sciences
ISSN: 1308 7231 (NWSAENS)
ID: 2018.13.1.1A0398

Status : Original Study
Received: October 2017
Accepted: January 2018

Funda Kutlu Onay
Amasya University, funda.kutlu@amasya.edu.tr, Amasya-Turkey
Cemal Köse
Karadeniz Technical University, ckose@ktu.edu.tr, Trabzon-Turkey

DOI	http://dx.doi.org/10.12739/NWSA.2018.13.1.1A0398	
ORCID ID	0000-0002-8531-4054	0000-0002-5982-4771
CORRESPONDING AUTHOR	Funda Kutlu Onay	

CLASSIFICATION OF FUNCTIONAL NEAR-INFRARED IMAGING BASED HEMODYNAMIC PATTERNS RECORDED AT MENTAL ARITHMETİC AND RESTING

ABSTRACT

Functional near-infrared spectroscopy (fNIRS) is a non-invasive optical imaging technique used in brain-computer interface (BCI) systems. It is used to measure deoxyhemoglobin and oxyhemoglobin proportions that occur during a specific activity in the brain region (motor and visual activity, auditory stimulus, etc.). In this study, hemodynamic patterns were recorded from 8 participants during mental arithmetic and rest activities. Features have been extracted for this by using detrended fluctuation analysis, entropy and Hjorth parameters methods. The distinctive feature vectors obtained after the feature selection process have been applied to support vector machines (SVM), multilayer artificial neural networks (MLANN) and k-nearest neighbors (k-NN) classifiers. As a result, the best classification accuracy was 97.17% when SVM classifier was used.

Keywords: Functional Near-Infrared Spectroscopy, Brain-Computer Interface, Detrended Fluctuation Analysis, Hjorth Parameters, Support Vector Machines

1. INTRODUCTION

Since the early 90's, cognitive studies have been performed using optical brain imaging systems. There are studies on human brain functions with brain imaging systems such as functional near-infrared spectroscopy (fNIRS) and positron emission tomography (PET) [1]. fNIRS is an optical imaging system that is usually used to understand brain regions and their relationship to mental functions and to measure brain functions for this purpose. Rather, there are studies that are measured during cognitive, visual, visuo-motor and motor tasks in humans [2]. It is desired to determine changes in hemoglobin (Hb) concentration of brain tissues. According to this method, brain functions related to non-invasive oxygenation, changes in blood oxygenation and blood density are displayed by viewing the blood flow in the forebrain [3]. Metabolic responses arising as a result of each task, such as oxy- and deoxyhemoglobin changes during a point-of-focus, allow the using of fNIRS or hybridization of electroencephalogram (EEG) and fNIRS in BCI systems. Accordingly, compared with EEG as another non-invasive method, NIRS-based BCI systems have several advantages. Due to no necessity for electrode gel, lack of electrooculographic artifacts arising from eye movements, practical sensor placement and become user-friendly, it is a preferable system for daily applications [4]. In this study, we performed classifying hemodynamic patterns which are recorded during mental arithmetic and resting. Features were extracted from cases

How to Cite:
Kutlu Onay, F. and Köse, C., (2017). Classification of Functional Near-Infrared Imaging Based Hemodynamic Patterns Recorded at Mental Arithmetic and Resting, **Engineering Sciences (NWSAENS)**, **13**(1):27-36, DOI: 10.12739/NWSA.2018.13.1.1A0398.

using detrended fluctuation analysis (DFA), entropy and Hjorth parameters. The chi-square and information gain tests were used with the identifying of selective features. Following the selection of the distinctive features, classification with maximum accuracy was provided with the appropriate classifier.

2. RESEARCH SIGNIFICANCE

In this study, the records obtained with non-invasive optical imaging technique which is called fNIRS was used. With this technique, changes in the concentration of non-oxygenated hemoglobin (deoxyhemoglobin) and oxygenated hemoglobin (oxyhemoglobin) can be displayed when an external stimulus is applied or during motor imaging task. This allows fNIRS technology to be used in BCI systems. The hemodynamic patterns of subjects recorded during resting and mental arithmetic are represented by the extracted features (DFA, entropy and Hjorth descriptors) and presented to the classifiers (MLANN, k-NN and SVM). Accordingly, the highest classification success was achieved when SVM was used.

3. MATERIALS AND METHODS

3.1. Dataset Description

The studied dataset consists of hemodynamic response samples during mental arithmetic (MA) tasks acquired from eight participants (three males, five females, mean age 26 years and SD 2.8 years) [5]. Participants performed cue-based mental arithmetic calculations. As a task, they were asked to subtract a single-bit number from a two-bit number. The numbers to be subtracted were displayed on the screen for 12 seconds followed by 28 seconds of resting. In this way, the participants performed 18 or 24 trials.

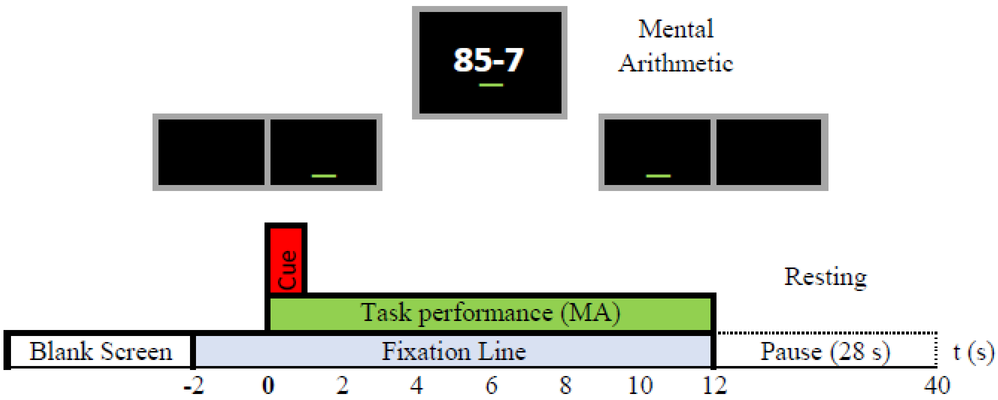

Figure 1. Timeline of one trial

3.2. Dataset Recording System

A continuous wave system (ETG-4000, Hitachi Medical Co., Japan) was used to acquire brain oxygenation. Oxy-, deoxy- and total-hemoglobin (Hb) changes is measured with a multi-channel recording system which has 16 photo-detectors and 17 light-emitters (3x11 grid, 52 channels). The sampling rate was set to 10 Hz and the distance between source and detector is 3cm. The figure of the recording system is modified from [6] and given in Figure 2.

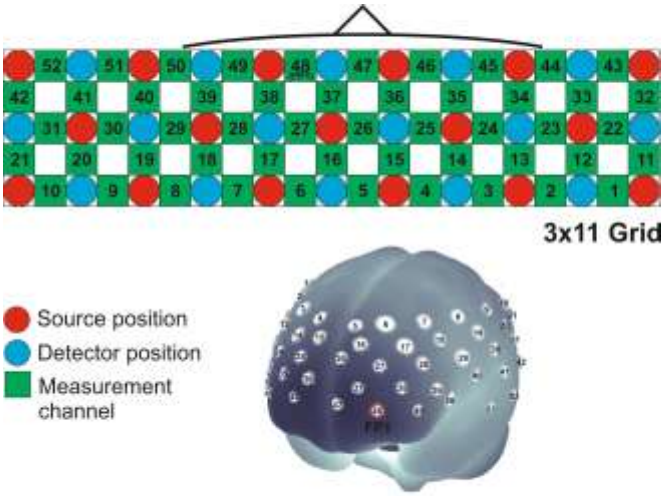

Figure 2. Figural representation of the recording system

4. FEATURE EXTRACTION AND SELECTION

4.1. Detrended Fluctuation Analysis (DFA)

DFA is used to characterize the correlation and similarity properties of non-stationary time signals. It was proposed by Peng et al. in 1994 [7]. In addition to conventional methods such as spectral analysis and Hurst analysis, DFA can retain intrinsic self-similarity embedded in apparently nonstationary time series and prevents misperception of self-similarity. DFA is also used in the analysis of biological data as well as in areas such as DNA, heart rate dynamics, long-term weather forecasting, economic time series, and solid state dynamics. In recent years, it has been successfully applied to a wide range of simulated and physiologic time series [8]. DFA analysis can be explained in 5 steps [9].

- The time series signal x(i), i=1. L, and L is the length of the signal. The cumulative sum of x(i) occurs as y(i).

$$y(i)=\sum_{j=1}^{i}(x(j)-\bar{x}),\ i=1,...,L \tag{1}$$

$$\bar{x}=\frac{1}{L}\sum_{j=1}^{L}x(j) \tag{2}$$

- The integrated time series y(i) is separated into m parts which m pieces of equal length l (L=ml). For each part, local trends are calculated with least square errors (LSE) method. For the detrending of y(i), local trends are subtracted.

$$Y_l(i)=y(i)-y_{k_l}(i),\ k=1,...,m \tag{3}$$

$$i=(k-1)l+1,...,kl \tag{4}$$

- For the kth part of length l, the root mean square (RMS) fluctuation for the integrated and detrended time series is calculated.

$$f_k(l)=\sqrt{\frac{1}{l}\sum_{j=(k-1)l+1}^{kl}\left[Y_l(j)\right]^2} \tag{5}$$

- Accordingly, the mean RMS fluctuation for entire time series is calculated as:

$$F(l)=\sqrt{\frac{1}{m}\sum_{k=1}^{m}f_k^2(l)}S \tag{6}$$

- Here β exponent is obtained when F(l) is computed for all l values.

$$F(l):l^{\beta} \tag{7}$$

4.2. Entropy Metrics

Entropy is a measure that gives the amount of information in a situation or variable. From a signal perspective, the entropy of a signal is a measure of the amount of information it contains; which corresponds to the distribution of the signal. The entropy of the data with a large and flat probability distribution is high. Conversely, the entropy of the data with a narrow and sharp distribution is low [10]. For EEG signals, entropy is another statistical descriptor that shows variability within. If the probability distribution function of a numerical random variable in the finite length is expressed by p(x), the calculation of the entropy of this variable becomes as in the equation (8). At the same time, this equation is called Shannon entropy [11].

$$H(x) = -\sum_{t=1}^{N-1} p_i(x) \log_2 (p_i(x)) \tag{8}$$

If x is a probabilistic experiment and *p(x)* is the probability of occurrence of this *x* result, then the average amount of information in this experiment is *H(X)*. Besides the Shannon entropy metric, approximate entropy is often used in studies.

4.3. Hjorth Descriptors

The Hjorth descriptors method is one of the nonlinear decomposition techniques in time series. Hjorth descriptors consist of the parameters of activity, mobility and complexity obtained in relation to the mean value and derivative of a signal. These were proposed by Hjorth in 1970 for the analysis of EEG signals [12].

The activity value is equal to the variance of the signal and is calculated by using the equation (9) of the *X* value of the sample of *N* in time space.

$$A(X_i) = \sigma_{X_i}{}^2 = \frac{\sum_{i=1}^{N} (X_i - \bar{X}_i)^2}{N-1} \tag{9}$$

Here, $\bar{X}$ shows the average of *X*.

The mobility parameter is a measure of the mean frequency of the signal and is calculated from the equation (10).

$$M(X_i) = \frac{\sigma_{\dot{X}_i}}{\sigma_{X_i}} = \frac{\sqrt{\frac{\sum_{i=1}^{N} (\dot{X}_i - \bar{\dot{X}}_i)^2}{N-1}}}{\sqrt{A(X_i)}} \tag{10}$$

Here, $\dot{X}$ corresponds to the first derivative of *X*.

The complexity parameter is a measure of the deviation of the signal from the sinus signal, and is obtained using equation (11).

$$C(X_i) = \frac{\sigma_{\ddot{X}_i} / \sigma_{\dot{X}_i}}{\sigma_{\dot{X}_i} / \sigma_{X_i}} \tag{11}$$

$\ddot{X}$ represents the second derivative of *X*.

4.4. Chi-Square Algorithm for Feature Selection

The Chi-square hypothesis test is used to test whether the two variables are dependent or independent. The goal of the feature selection based Chi-square test is calculation of dependency of a pattern within the class. While high Chi-square value is more descriptive for a pattern, zero value is indicating that the pattern is independent within this class. Chi-square value is calculated by the following equation (12) [13].

$$\chi^2 = \sum_{i=1}^{2}\sum_{j=1}^{k} \frac{(A_{ij} - E_{ij})}{E_{ij}} \quad (12)$$

In this equation; k, number of classes, A_{ij}, observed frequency value (i, rows; j, columns) and E_{ij}, expected (theoretical) frequency value.

5. CLASSIFICATION

5.1. Support Vector Machines (SVM)

SVM is a classification method based on supervised learning that is used in both linear and non-linear classification problems. Basically, it is based on the determination of decision boundaries that is hyperplanes, which separate the data belonging to the two classes most appropriately [14]. Classification of two classes that can be linearly separated is the most basic classification problem for SVM. There may be more than one hyperplane separating the two classes. However, the main purpose of support vector machines is to find a hyperplane that maximizes the distance between the nearest points. The optimal hyperplane which makes the most appropriate distinction by increasing the decision limit to the maximum and the points that line of this boundary width are called support vectors [15]. For most real-time applications, linear decomposition is not possible. The linear separable model of the SVM is also the basis for the nonlinear model. A non-linearly separable data in input space is defined in another space of higher dimension than the input space. This space is also called feature space [16]. A sample for a non-separation problem is shown in Figure 3. This transformation is performed by functions called kernel functions. Thus, the data can be linearly separated and hyperplane can be determined. The corresponding mathematical expression is given in the equation (13).

$$f(x) = \mathrm{sign}\left(\sum_i a_i y_i \varphi(x).\varphi(x_i) + b\right) \quad (13)$$

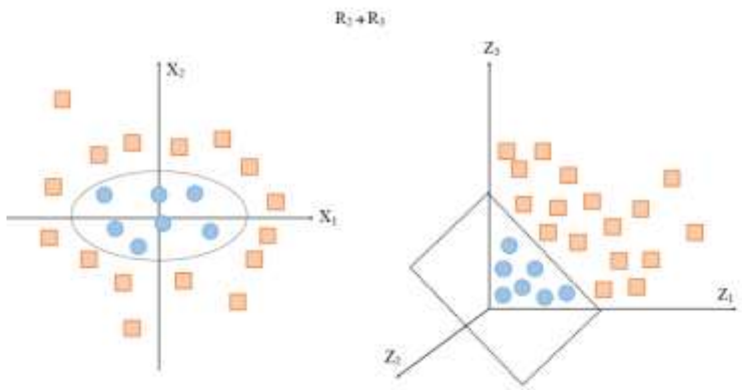

Figure 3. Linearization of a linearly non-separable problem with kernel-SVM

In linear applications, input data is given to the system by applying inner multiplication, while in complex applications; input data is mapped to the space by passing through a kernel function. φ in the equation (14) represents the kernel function. Determining the kernel functions to be used in SVM applications and the optimal parameters for them are important steps. Polynomial, radial basis function, Pearson VII function and normalized polynomial kernels, which are the most commonly used kernel functions in the literature, are given in Table 1 with the formulas and parameters.

Table 1. Kernel functions for SVM

Kernel Function	Mathematical Expression	Parameters
Polynomial kernels	$K(x,y) = ((x.y+1)^d$	Polynomial degree (d)
Normalized polynomial kernels	$K(x,y) = \dfrac{((x.y+1)^d}{\sqrt{((x.x)+1)^d((y.y)+1)^d}}$	Polynomial degree (d)
Pearson VII function	$\dfrac{1}{\left[1+\left(\dfrac{2\sqrt{\lvert x-y\rvert^2}\sqrt{2^{1/\omega}-1}}{\sigma}\right)^2\right]^{\omega}}$	Pearson width (σ,ω)
Radial basis kernel function	$K(x,y) = e^{-\gamma\lvert x-x_i\rvert^2}$	Kernel size (γ)

5.2. Multilayer Artificial Neural Networks (MLANN)

Multilayer Artificial Neural Networks (MLANN) is a neural network model consisting of one or more layers. Input data are taken from the input layer and distributed to the neurons in the hidden layer(s). The output layer receives the information from the hidden layer(s) and generates the output value. The number of layers and neurons in the hidden layer(s) are important parameters in the best separation [17].

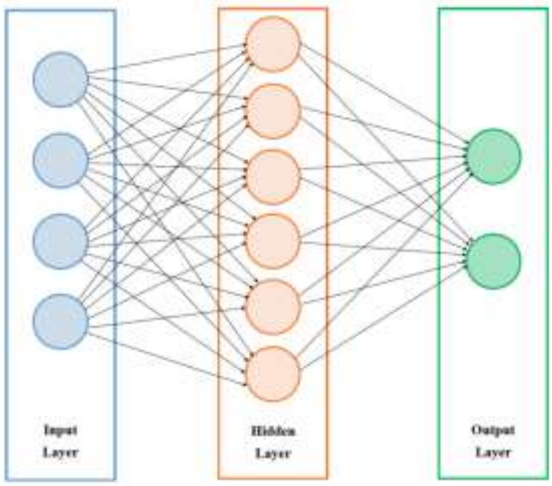

Figure 4. Architecture of a multilayer artificial neural network (MLANN)

In MLANN, training samples are presented to the network one by one and the calculation of the weight values is carried out which can produce the output within the error rate initially set.

5.3. k-Nearest Neighbors (k-NN)

The k-NN algorithm is a supervised learning method used in classifying data in applications related to pattern recognition and automatic learning [18]. According to this method, the k nearest neighbor of the pattern to be recognized is determined. Then, it is determined which class belongs to these k neighbors, and the label value of that class is assigned to the pattern sample to be recognized.

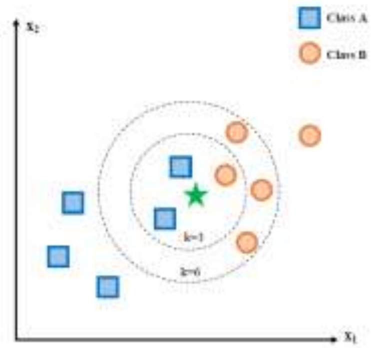

Figure 5. A sample for k-NN classifier

The k parameter used here is an important parameter that affects the classification accuracy. Various methods have been proposed in the selection of the k parameter. Due to the ease of implementation, trial-and-error or cross-validation methods are generally used. Distance calculation metrics which calculates similarity or distance between two points and there are many distance calculation methods in the literature (Euclidian, Mahalanobis, Manhattan, Chebshev etc.). However, in general, when calculating the distance between two points, the most commonly used method is the Euclidean distance method. The distance between two points, such as A and B, assumed to be N dimensional in space, is calculated by equation (14).

$$\mathrm{d} = \sqrt{(\mathrm{A}_1 - \mathrm{B}_1)^2 + (\mathrm{A}_2 - \mathrm{B}_2)^2 + \cdots + (\mathrm{A}_\mathrm{N} - \mathrm{B}_\mathrm{N})^2} \qquad (14)$$

6. RESULTS AND DISCUSSION

In this study, DFA, entropy and Hjorth descriptors are applied for classifying hemodynamic patterns which are recorded during mental arithmetic and resting. At the end of this, 12 features are obtained for all 6 cases. The recorded cases and explanations are given in Table 2.

Table 2. Obtained cases and descriptions

Cases	Description
C1	Oxygenation during mental arithmetic (oxy-MA)
C2	Oxygenation during resting (oxy-R)
C3	Deoxygenation during mental arithmetic (deoxy-MA)
C4	Deoxygenation during resting (deoxy-R)
C5	Total hemoglobin during mental arithmetic (total-MA)
C6	Total hemoglobin during resting (total-R)

After feature extraction step, the Chi-square test is applied to determine the most selective features for classes. As a result, the Chi value for all attributes was 0.9988 out of 1, and all the features were included in the classification process. MLANN, SVM and k-NN classifiers are applied to this feature vectors. The compared cases and classification results of these classifiers are given in Table 4, 5 and 6. Table 4 presents the SVM classification results along with the standard deviation and optimum σ values. In this study, radial basis kernel function was used and selection of the σ parameter directly affects the classification accuracy. The mathematical expression of RBF is expressed as (15). The best σ value was determined by 10-fold cross validation (10-FCV) and classification results were obtained.

$$K(x,y)=e^{-g|x-x_i|^2} \tag{15}$$

As a result of the implementation of the SVM classifier, the best classification accuracy was obtained in the C1-C3 case at 97.17%. This corresponds to the classification of oxy-Hb and deoxy-Hb changes in mental arithmetic. Similarly, we also obtained a considerable classification accuracy for oxy-Hb and deoxy-Hb changes at rest.

Table 3. Classification accuracy of SVM classifier
(CA: Classification accuracy, SD: Standard deviation)

Compared Cases	CA (%)	SD	Best σ
C1-C2	94.78	0.26	0.7
C3-C4	87.87	0.16	0.5
C5-C6	96.6	0.21	1.1
C1-C3	**97.17**	0.15	0.9
C2-C4	**97.04**	0.11	1.1
(C1,C3,C5)-(C2,C4,C6)	91.35	0.13	0.7

Table 4. Classification accuracy of MLANN classifier
(CA: Classification accuracy, SD: Standard deviation)

Compared Cases	CA (%)	SD
C1-C2	91.37	0.26
C3-C4	84.64	0.57
C5-C6	**96.82**	0.47
C1-C3	95.23	0.56
C2-C4	95.55	0.27
(C1,C3,C5)-(C2,C4,C6)	83.98	0.92

A MLANN model comprising 12 neurons in the input layer, 40 neurons in the hidden layer and 1 neuron in the output layer is designed. The MLANN classification results obtained are given in Table 4. According to this, MLANN is the best classification accuracy when the total-Hb changes in mental arithmetic and resting are classified.

Table 5. Classification accuracy of k-NN classifier
(CA: Classification accuracy, SD: Standard deviation)

Compared Cases	CA (%)	SD	Best k
C1-C2	80.75	0.07	22
C3-C4	83.84	0.01	23
C5-C6	85.38	0.23	5
C1-C3	85.03	0.05	20
C2-C4	80.15	0.04	25
(C1,C3,C5)-(C2,C4,C6)	**90.43**	0.11	6

The determination of the best k value for k-NN is very important for classification performance. For this, as in the SVM classifier, the best k values were determined with 10-FCV for all cases and shown in Table 5. This shows that the k-NN method gives the best result in classifying mental arithmetic and resting states. If we look at the general situation, it can be said that the most successful classification method in this study is SVM within the determined parameters. As a result, a higher performance was achieved in the classification performance of [5].

NOTICE

The study had been previously presented in the I. International Scientific and Vocational Studies Congress (BILMES 2017) in Ürgüp.

REFERENCES

1. Son, I.Y. and Yazici, B., (2006). Near Infrared Imaging and Spectroscopy for Brain Activity Monitoring. In Advances in Sensing with Security Applications, 341-372, Dordrecht, Springer.
2. Herff, C., Heger, D., Putze, F., Hennrich, J., Fortmann, O., and Schultz, T., (2013). Classification of Mental Tasks in the Prefrontal Cortex Using fNIRS. In Engineering in Medicine and Biology Society (EMBC) 35th Annual International Conference of the IEEE, 2160-2163.
3. Kontos, A.P., Huppert, T.J., Beluk, N.H., Elbin, R.J., Henry, L.C., French, J., Dakan S.M., and Collins, M.W., (2014). Brain Activation During Neurocognitive Testing Using Functional Near-Infrared Spectroscopy in Patients Following Concussion Compared to Healthy Controls. Brain imaging and behavior, Volume: 8, Issue: 4, 621-634.
4. McKendrick, R., Parasuraman, R., and Ayaz, H., (2015). Wearable Functional near Infrared Spectroscopy (fNIRS) and Transcranial Direct Current Stimulation (tDCS): Expanding Vistas for Neurocognitive Augmentation. Frontiers in Systems Neuroscience, Volume 9.
5. Bauernfeind, G., Steyrl, D., Brunner, C., and Müller-Putz, G. R., (2014). Single Trial Classification of fNIRS-based Brain-Computer Interface Mental Arithmetic Data: a comparison between different classifiers. In Engineering in Medicine and Biology Society (EMBC) 36th Annual International Conference of the IEEE, 2004-2007.
6. Bauernfeind, G., (2012). Using Functional Near-Infrared Spectroscopy (fNIRS) for Optical Brain-Computer Interface (OBCI) Applications.
7. Hardstone, R., Poil, S.S., Schiavone, G., Jansen, R., Nikulin, V.V., Mansvelder, H.D., and Linkenkaer-Hansen, K., (2012). Detrended Fluctuation Analysis: a Scale-Free View on Neuronal Oscillations. Frontiers in physiology, Volume: 3.
8. Peng, C-K., Hausdorff, J.M., and Goldberger, A.L., (2000). Fractal Mechanisms in Neural Control: Human Heartbeat and Gait Dynamics in Health and Disease. Self-Organized Biological Dynamics and Nonlinear Control, Cambridge, Cambridge University Press.
9. Chen, X., Wong, S.C., Tse, C.K., and Trajković, L., (2009). Detrended Fluctuation Analysis of the TCP-red Algorithm. International Journal of Bifurcation and Chaos, Volume: 19, Issue: 12, 4237-4245.
10. Najarian, K., and Splinter, R., (2005). Biomedical Signal and Image Processing, CRC press.
11. Lesne, A., (2014). Shannon Entropy: A Rigorous Notion at the Crossroads between Probability, Information Theory, Dynamical Systems and Statistical Physics. Mathematical Structures in Computer Science, Volume: 24, Issue: 3.
12. Hjorth, B., (1975). Time Domain Descriptors and Their Relation to a Particular Model for Generation of EEG Activity. Computerized EEG Analysis: CEAN, 3-8.
13. Liu, H., and Motoda, H., (2007). Computational Methods of Feature Selection, CRC Press.

14. Burges, C.J., (1998). A Tutorial on Support Vector Machines for Pattern Recognition. Data Mining and Knowledge Discovery, Volume: 2, Issue: 2, 121-167.
15. Hastie, T., Tibshirani, R., and Friedman, J., (2009). Overview of Supervised Learning. In the Elements of Statistical Learning, 9-41, New York, Springer.
16. Amari, S.I. and Wu, S., (1999). Improving Support Vector Machine Classifiers by Modifying Kernel Functions. Neural Networks, Volume 12, Issue 6, 783-789.
17. Patterson, D.W., (1998). Artificial Neural Networks: Theory and Applications, Prentice Hall PTR.
18. Kotsiantis, S.B., Zaharakis, I., and Pintelas, P., (2007). Supervised Machine Learning: A Review of Classification Techniques, IOS Press.