

EEG Signal Classification using Neural Network and Support Vector Machine in Brain Computer Interface

M M El Bahy¹, M.Hosny¹, Wael A. Mohamed¹, Shawky Ibrahim¹

¹ Faculty of Engineering (at Benha), Benha University, Cairo, Egypt,
{mmbahy, mohamed.hosny, wael.ahmed, shawky.Ibrahim}@bhit.bu.edu.eg

Abstract. Classification of EEG signals is one of the biggest problems in Brain Computer Interface (BCI) systems. This paper presents a BCI system based on using the EEG signals associated with five mental tasks (baseline, math, mental letter composing, geometric figure rotation and visual counting). EEG data for these five cognitive tasks from one subject were taken from the Colorado University database. Wavelet Transform (WT), Fast Fourier Transform (FFT) and Principal Component Analysis (PCA) were used for features extraction. Artificial Neural Network (ANN) trained by a standard back propagation algorithm and Support Vector Machines (SVMs) were used for classifying different combinations mental tasks. Experimental results show the classification accuracies achieved with the three used feature extraction techniques and the two classification techniques.

Keywords: Brain Computer Interface (BCI), Artificial Neural Network (ANN), Support Vector Machine (SVM)

1 Introduction

Today, a lot of communication systems between disabled peoples and external devices have been suggested and developed. Among them, brain computer interface (BCI) and brain machine interface (BMI) has been very attractive recently. One mental task can be visualized by a person then the brain wave signals are measured, processed and evaluated to identify this mental task. Moreover, external devices such as computers or machines can be controlled [1].

AT first BCIs were used for medical reasons, but now BCI systems are also being developed for general people purposes mostly for entertainment. The main technology used in BCI systems for recording brain activity is electroencephalography (EEG). Although it is an imperfect indication of brain activity, compared to other technologies (MEG, FMRI and FNIR), but EEG has the most advantages for BCI systems. The main advantages are high temporal resolution, portability, doesn't expose patients to high-intensity magnetic fields and low cost of EEG hardware [2].

We can divide BCI systems in to the following classes: invasive BCI in which sensors have surgically implanted into the brain and non-invasive BCI that uses sensors located on the scalp [3]. The use of additional motor movements is required in dependent BCI, while independent BCI doesn't require any muscle activity. A synchronous BCI where user interacts with the system only in specific time frames, however asynchronous can be used for any time frames.

BCI systems methodology consist of signal acquisition, preprocessing, feature extraction and classification. So far the accuracy of classification has been one of the main drawbacks of the current BCI systems. Enhancing the accuracy may be achieved through enhancements in the three main phases of BCI.

This paper presented a non-invasive offline system for classifying different combination of mental tasks using EEG signals from the publicly available dataset of Kein and Aunon's and it achieved a high classification rates. In this research three different techniques of features extraction were used which are: Wavelet Transform, Fast Fourier Transform and Principal Component Analysis. Two classifications techniques were used which are: Neural Network trained by a standard back propagation algorithm and Support Vector Machines.

The rest of this paper continues as follows. Section 2 presents the previous work for brain computer interface systems. Section 3 describes the system methodology and explains the used techniques for feature extraction and data classification. Section 4 illustrates the experimental results of the proposed methodology for classifying the mental tasks. Conclusion and future work are illustrated in section 5.

2 RELATED WORK

N. Saadat, and H. Pourghassem [4] acquired EEG signals from three normal subjects during three tasks: Imagination of the left, right hand movements and generating some words. Band pass filter between 0-

30HZ was used to remove noise from EEG signals and the transition matrix of EEG signals was scaled between [0, 1] using non-linear normalization technique. Discrete Fourier Transform (DFT) was used to extract the spectral and spatial features from EEG signals as a feature extraction method. Classification was done using multi-layer perceptron neural network trained by back propagation algorithm. Classification rates between 73% and 81% were achieved for all subjects.

Kenji Nakayama et al [5] presented efficient pre-processing techniques in order to achieve high classification accuracy of mental tasks. The preprocessing techniques like segmentation along time axis, amplitude of FFT of EEG signals and reduction of samples by averaging and nonlinear normalization. Classification accuracy of 78 % was achieved for the recognition of five tasks.

Anderson et al. [6] proposed a system by which EEG features were extracted through the short time principal component analysis (STPCA) and the EEG data was classified by the linear discriminant analysis (LDA). Classification accuracy for the recognition of five tasks was 77.9 %.

Yuji Mizuno et al. [7] employs the maximum entropy method (MEM) for frequency analyses and investigates an alpha frequency band and beta frequency band in which features are more apparent. In addition, learning vector quantization (LVQ) is used for clustering the EEG data with features extracted and classification accuracy for the recognition of five tasks was 81 %.

Hosni et al [8] used three of the five mental tasks from Keirn and Aunon's dataset. These tasks were baseline task, letter task and math task. Eye blinks were identified and removed using Independent Component Analysis (ICA). Three different feature extraction techniques were used in this paper which are Parametric Auto Regressive (AR) modeling, AR spectral analysis and band power differences. Classification was done using Radial Basis Function (RBF) and Support Vector Machines (SVM). Best classification accuracy achieved was 70%.

Martina Tolić and Franjo Jović [9] extracted the features of EEG signals using Discrete Wavelet Transform and Neural Network was used as a classifier. Mean classification accuracy for the recognition of all five tasks was 90.75%.

3 PROPOSED SYSTEM DESIGN

The aim of this research is to compare between three different features extraction techniques with two classifiers and classifying different combinations of three, four and five mental tasks. The system's methodology comprises of four main stages as illustrated in Fig. 1. The first step was Signal acquisition. The second step was signal preprocessing, to remove noises, artifacts and unwanted data. The third step was features extraction from the EEG signals. The fourth and the final step was classification of the signals to different classes that corresponding to the different mental tasks.

3.1 Signal acquisition

The EEG data used in this study were collected by Keirn and Aunon [10]. This dataset can be described as follows:

Several trials of five mental tasks were recorded and the number of times that each mental task was repeated is different from one subject to another. The number of trials for each subject as shown in Table 1[11].

Each channel from each trial produced 2500 sample points for the 10 s recording because the amplified EEG signals were sampled at 250 samples per second. The selected subject was 6.

EEG signals were recorded in two different days so, there were two sessions of recordings. First session differs from second session, this is possible since the recording sessions were separated by two weeks and it is known that the statistics of the brain waves are non-stationary over extended periods of time. If the subject is losing concentration throughout the task then mixing early and late portions of the EEG may degrade classification performance. In our research EEG signals of both sessions were used together and this is one of the main challenges in our research.

Data for seven subjects were recorded, that every subject was seated in an industrial acoustics company sound controlled booth with dim lighting and a noiseless fan. EEG signals were recorded from positions C3, C4, P3, P4, O1 and O2 (shown in Fig. 2) using an electro-cap (elastic electrode cap) defined by the 10–20 system of electrode placement [12].

The impedances of all electrodes were retained below 5 kilo ohm. They made measurements with reference to electrically linked mastoids, A1 and A2. A bank of amplifiers (Grass7P511) were connected through the electrodes whose band-pass filter were set at 0.1 to 100 Hz to preprocess the data. The EEG signals were sampled at 250 Hz with a Lab Master 12-bit A/D converter mounted on a computer.

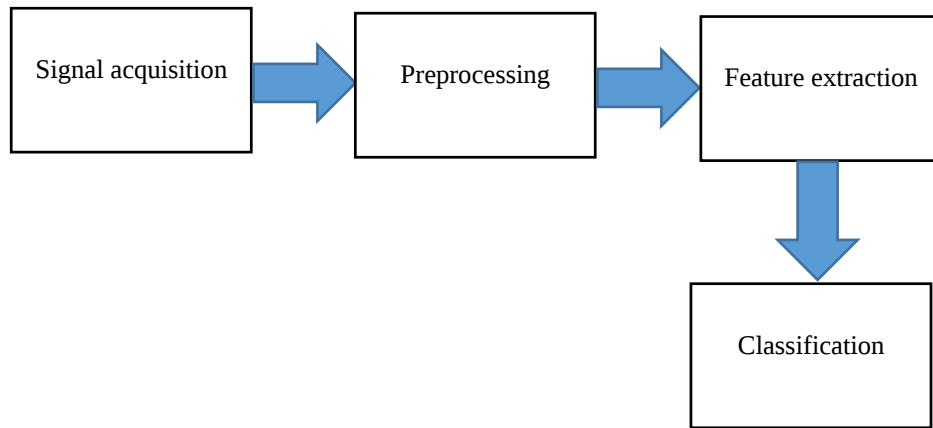


Fig. 1. Proposed system methodology.

Table 1. No of trials for each subject

Subject	Number of trials
1	10
2	5
3	10
4	10
5	15
6	10
7	5

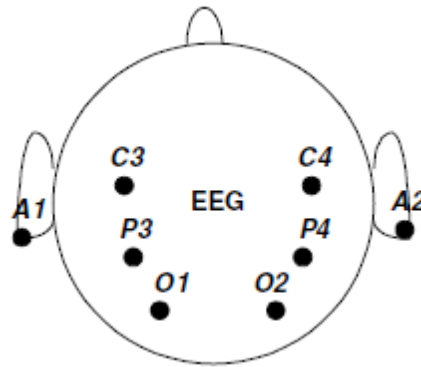


Fig. 2. Electrodes placement.

In this paper, EEG signals from subject 6 performing five different mental tasks have been used. These mental tasks are:

Baseline task. Every subject was asked to relax and think of nothing in particular. This task can be used as a control and as a baseline measure of the EEG signals.

Math task. The subjects were given none trivial multiplication problems, such as 45 times 18 and they were supposed to solve them without vocalizing or making any movements. The task was none repeating and designed so that an immediate answer was not apparent. At the end of the task the subjects verified whether or not he/she attained the solution and no subject finished the task before the end of the 10 s recording session.

Mental letter-composing task. Every subject was asked to mentally compose a letter to a close friend without vocalizing. This task was done several times and every subject was asked to continue with the letter from where they left off the previous time rather than starting again each time.

Geometric figure rotation task. Every subject was given 30 s to study a particular three-dimensional object, after which the drawing was removed and the subjects were asked to visualize that object being rotated about an axis.

Visual counting task. Every subject was asked to imagine a blackboard and to visualize numbers being written on the board consecutively, with the previous number being erased before the next number was written. They were also told to resume counting from where they left off in the previous time rather than starting again each time.

3.2 Preprocessing

In preprocessing stage noise and artifacts should be removed to enhance classification accuracy so, band pass filter between 1 and 45Hz was used to filter the EEG signals and a 5th order Butterworth filter was used to remove the unwanted artifacts.

Band pass filter helps to select the frequency band containing useful information, reducing the number of features used for classification, have a direct influence on reducing the execution time of the system, and increasing the utilization of Memory which improve the system performance.

3.3 Feature extraction

Three different features extraction techniques which are wavelet packet decomposition, fast Fourier transform and principal component analysis were implemented to compare between their performances with two classifiers.

Many methods such as time domain, frequency domain, and time-frequency domain methods were used [13]. Wavelet transform (WT) is considered to be one of the most suitable choice to use time-frequency domain methods for feature extraction so it was the first used feature extraction technique [14]. The output of the Wavelet packet decomposition can be computed by the following equation:

$$Wp_t = wpdec(x, Level, 'haar') \quad (1)$$

Where $wpdec$ is a one-dimensional wavelet packet analysis function and $Level$ split the data vector x into tree nodes for making the computation in each node. The EEG data were decomposed into (Haar) mother wavelet with five level wavelet packet decomposition in our system.

The second used feature extraction technique was the fast Fourier transformation, to extract the frequency components of the signal, select the required components and calculate the power for these components which were considered to be the input vector of each classifier. FFT computes the DFT where:

$$X_k = \sum_{n=0}^{N-1} x_n e^{-i2\pi k n / N} \quad (2)$$

Where $k=0, 1 \dots N-1$, X_n is the sampled values and N is the total number of samples in the vector [15].

The last used method was the Principal Component Analysis technique that generally used to dimensionally reduce the original data to first n Eigen values [16]. Principal component analysis is a statistical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components. Principal components number is less than or equal to original variables number.

In this transformation the first principal component has the largest possible variance value. The PCA transformation matrix $W = [e.g. w_1, w_2, \dots, w_n]$ can be obtained by performing a general eigenvalue decomposition of the covariance matrix $R = XX^t$ where X is the input signal(s) and $w_1, \dots, w_n$ are n normalized orthogonal eigenvectors of X^tX corresponding to n different eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ in descending order.

The PCA transformation (Y) of X is then given by: $Y = W^tX$ where the rows of Y are uncorrelated to each other.

3.4 Classification

Neural Network. Several researchers have been used neural network to classify the EEG signal. In this research artificial neural network trained by a standard back propagation algorithm was used for classification.

Data were recorded from seven subjects during performing five mental tasks which were (baseline, math, mental letter composing, geometric figure rotation and visual counting). The selected subject was 6. The five mental tasks were measured 10 times each one. The length of EEG signals for each trial was 10sec. The recorded data were divided into training and testing sets Therefore, 10 trials are available for each task. Among them, 9 trials are used for training and one trial is used for testing.

Ten trials are selected for testing and classification accuracy was evaluated based on the average over these 10 trials.

Extracted features were considered as input neurons to the neural network algorithm. The output layer should contain 5 neurons for the five classes that represent the five mental tasks that we want to classify. The number of neurons in the input layer varied according to the length of the input features vector.

Support vector machine (SVM). Support vector machine is a supervised learning method to analyze data and distinguish patterns, frequently used for classification and regression analysis.

SVM constructs a discriminant hyperplane that maximizes the margins to identify classes, compared with other classifiers; SVM has a good generalization property, insensitive to overtraining and has a good performance with limited data.

The five mental tasks were measured 10 times each one. The length of EEG signals for each trial was 10sec. The recorded data were divided into training and testing sets Therefore, 10 trials are available for each task. Among them, 9 trials are used for training and one trial is used for testing.

Ten trials are selected for testing and classification accuracy was evaluated based on the average over these 10 trials.

4 EXPERIMENTAL RESULTS

EEG data of subject 6, which can be obtained from the web site of Department of computer science, Colorado state university, were used. Three different features extraction methods were used as follows.

4.1 Wavelet Transform (WT):

Wavelet Packet Decomposition was applied on the EEG signals. The EEG data were decomposed into Haar mother wavelet with five level wavelet packet decomposition. Coefficients from nodes (5 0), (5 1), (5 2), (5 3), (5 4) and (5 5), which represents frequencies from 1 Hz to 45 Hz were extracted. The mean, $\mu(x)$, standard deviation, $\sigma(x)$, and entropy, $\varepsilon(x)$ were calculated for these coefficients by the following equations respectively:

$$\mu = \frac{1}{N} \sum_{i=1}^N x_i \quad (3)$$

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2} \quad (4)$$

$$\varepsilon = - \sum_{i=1}^N P(x_i) \log_2(x_i) \quad (5)$$

Where N is the total number Coefficients in the vector, P is the probability of x_i Values of each coefficient vector were calculated and used as features. Thus we have $6*3=18$ features for each channel and a total of 108 features for each task.

4.2 Fast Fourier Transform (FFT):

Spectrum of EEG signals was calculated and the top hundred Fast Fourier Transform power values were taken. Thus we have 100 features for each channel and a total of 600 features for the 6 channels used.

4.3 Principal Component Analysis (PCA):

PCA technique reduced the original data to first n Eigen values. The highest variance value was taken as a feature so we have one feature for each channel and a total of 6 features for the 6 channels.

Neural Network trained by a standard back propagation algorithm and support vector machine were used for classification.

4.4 Multi-layer Perceptron Neural Network

Neural Network trained by a standard back propagation algorithm was used in our research. The number of neurons in the input layer varied according to the length of the input features vectors. Many tests were done to find the best configuration for the neural network in terms of: number of neurons in the hidden layer and the maximum number of iterations (epochs) in the learning process.

For each features set, the configuration that produced optimal weights (which lead to maximum correct classification rate in the testing) for I/O mapping was used which were:

Number of neurons in the hidden layer=100.

Maximum number of iterations (epochs) in the learning process=1000.

The activation function used was the sigmoid function, the learning rate was 0.1 and the training stopped when either the maximum number of epochs reached 1000 or the mean square error reached to a small value such as 0.001.

4.5 Support vector machine

The SVM classifier in this paper was based on LIBSVM implementation from [17]. Many tests were done to find the optimal parameters for SVM in terms of: type of the kernel, the Coefficient in kernel function, Degree in kernel function. Parameters that lead to maximum correct classification rate in the testing for I/O mapping were used which were:

Polynomial kernel was used.

Degree in kernel function=3.

Coefficient in kernel function=0.

Data were analyzed using MATLAB 2013 and a computer (Intel Core i7 CPU 2.20 GHz, 8 GB DDR RAM, Windows 7).

Total classification accuracies for classifying different combination of three mental tasks using the three feature extraction techniques and two classifiers as shown in Table 2 and Table 3 shows the effect of increasing the frequency band from [1 45] to [1 100] on it.

Total classification accuracies for classifying different combination of four mental tasks as shown in Table 4 and Table 5 shows the effect of increasing the frequency band from [1 45] to [1 100] on it.

Table 6 shows classification accuracies for classifying all five mental tasks using frequency band [1 100].

Table 2. Classification accuracies of different three mental tasks, frequency band [1 45].

	B,m,l	B,m,r	B,m,c	B,l,r	B,l,c	B,r,c	M,l,r	M,l,c	L,r,c	M,r,c
Svm fft	70	90	70	80	56.67	83.34	80	53.33	63.333	83.335
Svmwav	86.668	96.67	86.67	90	76.67	93.33	96.67	73.33	76.668	86.67
Svm pca	73.335	80	83.335	86.668	70	90	66.667	56.666	73.335	73.336
Nn fft	66	82.33	66	67.67	46	78.33	71	50.33	64.67	76.33
Nn wav	90	96.67	77.67	90	77.67	86.67	93.33	80.67	83.33	83.33
Nn pca	75	76.67	80	86.67	78.67	86.67	73.33	72.67	79	70

BBase line task.

M....Math task.

L.....Letter composing task.

R.....Geometric figure rotation task.

C.... Visual counting task.

Table 3. Classification accuracies of different three mental tasks, frequency band [1 100].

	B,m,l	B,m,r	B,m,c	B,l,r	B,l,c	B,r,c	M,l,r	M,l,c	L,r,c	M,r,c
Svm fft	70	90	70	76.67	56.67	83.33	80	50	60	83.33
Svmwav	86.67	96.67	90	90	80	93.33	96.67	83.33	83.33	90
Svm pca	76.67	80	83.33	90	70	90	70	63.33	76.67	76.67
Nn fft	64.67	82.33	66	66.67	43.33	81.67	73.33	54.33	64.33	75
Nn wav	90	96.67	83.33	90	79	87.33	93.33	86	83.33	80
Nn pca	78	73.33	86.67	86.67	80.67	90	73.67	81.67	90	83.33

Table 4. Classification accuracies of different four mental tasks, frequency band [1 45].

	B,m,l,r	B,m,l,c	B,m,r,c	B,l,r,c	M,l,r,c
Svm fft	77.5	57.5	77.5	65	60
Svmwav	90	77.5	87.5	80	77.5
Svm pca	70	65	75	75	57.5
Nn fft	59	46.5	65.75	51.5	56.5
Nn wav	92.5	80	85.25	81.25	82.5
Nn pca	74	77.75	74.25	80	68.5

Table 5. Classification accuracies of different four mental tasks, frequency band [1 100].

	B,m,l,r	B,m,l,c	B,m,r,c	B,l,r,c	M,l,r,c
Svm fft	75	55	77.5	62.5	57.5
Svmwav	90	82.5	90	82.5	85
Svm pca	72.5	65	77.5	75	62.5
Nn fft	59.25	44.25	64.25	56	59.25
Nn wav	92.5	82.75	85	82.25	82.5
Nn pca	73.75	78.25	75	86	70.25

Table 6. Classification accuracies of all five mental tasks, frequency band [1 100].

Svm fft	60
Svmwav	84
Svm pca	64
Nn fft	45.2
Nn wav	82.6
Nn pca	74.2

As shown in the above tables the performance of wavelet transform is better than fast Fourier transform and principal component analysis whether with neural network or support vector machine. Performance of principal component analysis is better than fast Fourier transform with the two classifiers. Increasing the frequency band from [1 45] to [1 100] improves the classification accuracies. Best classification accuracy for classifying five mental tasks was 84% and it was obtained for wavelet packet decomposition with support vector machine.

A tri-state Morse code scheme could be used as an application of our system to help disabled peoples having problems in speech as it could translate different combination of three mental tasks into English words like food, water and TV. The basic alphabets in the conventional Morse code scheme are dot and dash so two mental tasks will be sufficient to be used but, the use of an additional mental task was proposed to represent space between dot and dash. This space will represent the end of either a dot or dash and starting of a new dot or dash, which help users to concentrate on the sequence of mental tasks not the time duration of each mental task. Therefore, to use this tri-state Morse code, we need three different combinations of mental tasks where each task will correspond to either a dot, a dash or a space. Using this tri-state Morse code, English letters, Arabic numerals and punctuation marks to form words and complete sentences could be constructed [18].

5 CONCLUSION

This paper presented a non-invasive system for classifying different combinations of mental tasks using Brain EEG signal processing. In the proposed model EEG data of subject 6 was obtained from the web site of Department of computer science, Colorado state university. The EEG signals were extracted by six electrodes (C3, C4, P3, P4, O1 and O2) for five mental tasks. Wavelet Transform (WT), Fast Fourier Transform (FFT) and Principal Component Analysis (PCA) techniques were used for features extraction. Data were classified using the Back Propagation neural network and support vector machine. Experimental results show the classification accuracies achieved with the three used feature extraction techniques and the two classification techniques.

REFERENCES

1. G. Dornhege, J. d. r. Millan, T. Hinterberger, D. J. McFarland and K. R. Muller.: Toward Brain-Computer Interfacing, The MIT Press, (2007)
2. Berger, T.W., Chapin, J.K., Gerhardt, G.A., McFarland, D.J., Principe, J.C., Soussou, W.V., Taylor, D.M., & Tresco.: P.A. Brain-computer interfaces: An international assessment of research and development trends. Springer, New York (2008)
3. E. Hortal : Online Classification of Two Mental Tasks Using a SVM based BCI System. IEEE Neural Engineering, (2013)
4. N. Saadat, and H. Pourghassem: Mental task classification based on HMM and BPNN. International Conference on Communication Systems and Network Technologies (CSNT), IEEE, (2013) 210-214
5. Kenji Nakayama and Kiyoto Inagaki: A Brain Computer Interface Based on Neural Network with Efficient Pre-Processing. International Symposium on Intelligent Signal Processing and Communication Systems, Yonago Convention Center Tottori Japan December (2006) 673-676
6. C.W.Anderson, and J A.Bratman: Translating thoughts in to actions by finding patterns in brain wave. In proceedings of the fourteenth Yale workshop on adaptive and learning systems, Yale University New Haven CT June (2008) 1-6
7. Yuji Mizuno and Hiroshi Mabuchi.: Clustering of EEG data using maximum entropy method and LVQ. International Journal of Computers issue 4, volume 4. (2010)
8. Hosni, M. Gadallah, S. Bahgat, and M. AbdelWahab.: Classification of EEG signals using different feature extraction techniques for mental-task bci. In Computer Engineering & Systems. ICCES'07 International Conference on IEEE, (2007) 220–226
9. Martina Tolić, Franjo Jović: Classification of wavelet transformed EEG signals with neural .network for Imagined mental and motor tasks. Kinesiology, (2013) 130-138
10. Z. A. Keirn, and J. I. Aunon: A New Mode of Communication between Man and His Surroundings. IEEE Transactions on Biomedical Engineering, vol. 37. (1990) 1209-1214
11. Nai-Jen Huan, Ramaswamy Palaniappan: Neural network classification of autoregressive features from electroencephalogram signals for brain–computer interface design. J. Neural Eng. 1, (2004) 142–150
12. Jasper H 1958 the ten twenty electrode system of the international federation Electroencephalographic Clin. Neurophysiol. **10** 371–5
13. H. Shedeed et al.: Brain EEG Signal Processing For Controlling a Robotic Arm. IEEE, (2013) 152-157
14. Guang Yang, Kenji Nakayama and Akihiro Hirano: A Dual-Class Voting Mechanism for Brain Computer Interface based on Wavelet Packet and Support Vector Machine. IEEE, (2013)
15. Howida A.Shedeed, Mohamed F.Issa: Brain-EEG signal classification based on data normalization for controlling a robotic arm. International journal of tomography and simulation, volume 29. (2016)
16. Kottaimalai R et al.: EEG Signal Classification using Principal Component Analysis with Neural Network in Brain Computer Interface Applications. IEEE International Conference on Emerging Trends in Computing, Communication and Nanotechnology (ICECCN), (2013)
17. Chih-Chung Chang and Chih-Jen Lin. LIBSVM: A library for support vector machines. ACM Transactions on Intelligent Systems and Technology, (2011) 2:27:1–27:27
18. Ramaswamy Palaniappan et al.: A New Brain–Computer Interface Design Using Fuzzy ARTMAP. IEEE transactions on neural systems and rehabilitation engineering, vol. 10. (2002)