Motor Imaginary Task Classification using Statistically Significant Time Domain and Frequency Domain EEG features

İstatistiksel olarak Anlamlı EEG Zaman Alanı ve Frekans Alanı Öznitelikleri ile Motor Hayali Görev Sınıflandırılması

Murside Degirmenci[1,*], Yilmaz Kemal Yuce[2], Yalcın Isler[3]
[1]Department of Biomedical Technologies, Izmir Katip Celebi University, Izmir, Turkey
[2]Department of Computer Engineering, Alanya Alaaddin Keykubat University, Antalya, Turkey
[3]Department of Biomedical Engineering, Izmir Katip Celebi University, Izmir, Turkey
ORCIDs: 0000-0003-0978-9653, 0000-0001-5291-0565, 0000-0002-2150-4756
E-mails: mrsddgrmnc.09@gmail.com, yilmazkemalyuce@gmail.com, islerya@yahoo.com
*Corresponding author.

***Abstract*—Motor Imaginary (MI) electroencephalography (EEG) signals are obtained when a subject imagines a task without essentially applying it. The accurate decoding of MI EEG signals plays an important role in the design of brain-computer interface (BCI) systems due to the use of these signals in the rehabilitation process of paralyzed patients in recent studies. In this study, two different MI tasks were tried to be differentiated by extracting time-domain and frequency-domain features from 22 channel EEG signals and determining best combination of important and distinctive features based on statistical significance. MI EEG signals were supplied from BCI Competition IV Dataset-IIa. These features were differentiated using 25 different classification algorithms and 5-fold cross-validation method. The repeatability of the results was examined testing each algorithm 10 times. As a result, the highest average accuracy rate of 60.69% was calculated in the Quadratic Support Vector Machine (SVM) using all features and 62.52% in the Ensemble Subspace Discriminant (ESD) algorithm using only the selected features by the independent t-test. The results showed that the independent t-test based feature selection increased the performance in 20 classifiers, and decreased the performance in 5 classifiers. Also, the effectiveness of the feature selection method examined using the paired-sample t-test which is known as repeated measures t-test. The significance value, p-value was found as 0.04. Therefore, the independent t-test based feature selection method is an effective feature selection method and is providing the significant improvement in classifier performance.**

Keywords*—EEG signals; feature selection; frequency-domain features; motor imaginary task classification; time-domain features

***Özetçe*—Motor Hayali (MH) elektroensefalografi (EEG) sinyalleri, bir özne, aslında uygulamadan bir görevi hayal ettiğinde elde edilir. Son yıllarda yapılan çalışmalarda bu sinyallerin felçli hastaların rehabilitasyon sürecinde kullanılmasından dolayı MH EEG sinyallerinin doğru olarak çözülmesi beyin-bilgisayar arayüzü (BBA) sistemlerinin tasarımında önemli bir rol oynamaktadır. Bu çalışmada, 22 kanallı EEG sinyallerinden zaman alanı ve frekans alanı öznitelikleri çıkarılarak ve istatistiksel anlamlılığa dayalı olarak önemli ve ayırt edici özelliklerin en iyi kombinasyonu belirlenerek iki farklı MI görevi ayırt edilmeye çalışılmıştır. MI EEG sinyalleri, BCI Competition IV Dataset-IIa'dan sağlandı. Bu öznitelikler, 25 farklı sınıflandırma algoritması ve 5-kat çapraz doğrulama yöntemi kullanılarak ayrıştırılmıştır. Sonuçların tekrarlanabilirliği, her bir algoritma 10 defa test edilerek incelenmiştir. Sonuç olarak, en yüksek ortalama doğruluk oranı tüm öznitelikler kullanılarak Kuadratik Destek Vektör Makinesinde (DVM) 60,69% ve sadece bağımsız t-testi ile seçilen öznitelikler kullanılarak Topluluk Altuzay Ayırım (TAA) algoritmasında 62,52% olarak hesaplanmıştır. Sonuçlar, bağımsız t-testine dayalı öznitelik seçiminin 20 sınıflandırıcıda performansı artırdığını, 5 sınıflandırıcıda ise performansı azalttığını göstermiştir. Ayrıca, yinelenen ölçümler t-testi olarak bilinen bağımlı örneklem t-testi kullanılarak öznitelik seçme yönteminin etkinliği incelenmiştir. Anlamlılık değeri olan p-değeri 0.04 olarak bulunmuştur. Bu nedenle, bağımsız t-testine dayalı öznitelik seçim yöntemi, etkili bir öznitelik seçme yöntemidir ve sınıflandırıcı performansında önemli iyileştirmeler sağlamaktadır.**

Anahtar Kelimeler*—EEG sinyalleri; öznitelik seçimi; frekans alanı öznitelikleri; motor hayali görev sınıflandırma; zaman alanı öznitelikleri

I. Introduction

The Brain computer interface (BCI) systems supplies an interface between the human brain and external device for converting the brainwaves of the subject into commands to

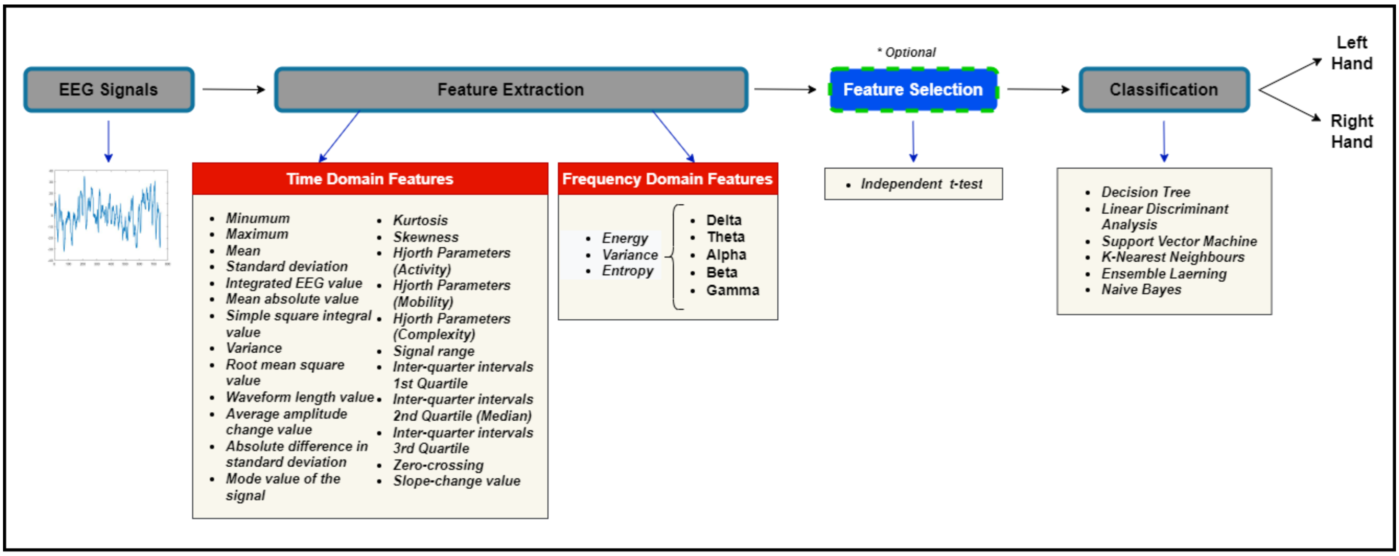

Figure 1: The flowchart of the proposed study

external device [1]. BCI systems mostly based on electroencephalography (EEG) signals which is electrical recording of brain activities obtained by placing electrodes on the scalp. The different types of EEG signals examined in research are categorized as P300 evoked potential [2], steady-state visual evoked potentials (SSVEPs) [3], and motor imaginary (MI).

MI EEG signals represent the neural activity that generates when a subject voluntarily imagines implementing a movement without actually performing it [1]. MI EEG signals have the low signal-to-noise ratio and the variation of brain waves between subjects, so these drawbacks make the processing of signals as difficult task. Therefore, EEG-based automatic detection algorithms have been started to be generated for the discrimination of motor imaginary tasks in recent decades.

The traditional machine learning (ML) algorithms proposed handcrafted feature extraction process in MI-BCI studies. MI features can be extracted using three different categories based on the domain in which data are processed: temporal features, spectral features and spatial features. Temporal features draw on time-domain using different time segments and time points: mean, variance, Hjorth parameters, skewness, mean absolute value, etc [4]. Spectral features consists of two main groups which are frequency-domain features (power spectral density (PSD) and fast Fourier transform (FFT) [5]) and time-frequency features (short-time Fourier transform (STFT), Wavelet transform (WT) [6] etc.). As spatial features, common spatial pattern (CSP) [7] and its different versions [8] such as filter bank common spatial pattern (FBCSP) [9] are mostly used algorithms in MI BCI research studies. The selection of effective features is one of the most important factors effecting algorithm performance [10]. If a large number of features are used together, the complexity of the classifier increases and in some cases the performance of the classifier decreases [11], [12]. In order to overcome this problem, different feature selection methods have been proposed in the literature in the analysis of physiological signals: statistical significance based selection, backward elimination, forward selection, principal component analysis (PCA), genetic algorithm (GA) [11]–[14].

Gaur et al. [15] proposed the subject specific multivariate empirical mode decomposition (MEMD) based filtering method to binary (left and right hand) classification of MI tasks and obtained accuracy of 79.93% using BCI Competition IV Dataset-IIa. In [16], the frequency bands were obtained based on WT and CSP calculated for four MI task classification. They found average accuracy of 64.4% using hierarchical support vector machine (HSVM) algorithm. In another study [17], the five-class MI task classification based on multi-class CSP was proposed to differentiate finger movements and classification accuracy calculated as 40.6%. Luo et al. [10] used Wavelet packet decomposition (WPD) to obtain frequency domain of EEG signals and features selected based on Dynamic frequency feature selection (DFFS) method for binary classification (left and right hand) of MI tasks. The average accuracy of 68.32% was calculated by Random Forest (RF) on BCI Competition IV Dataset-IIa.

In this study, a computer-aided approach is proposed to differentiate EEG-based MI tasks from each other. It is aimed to generate a successful classification algorithm extracting various time-domain and frequency-domain features from EEG signals. In this direction, it has been studied with 3 different spectral features extracted from 5 different frequency bands and 24 different temporal features. Also, the performance improvement of statistical significance based feature selection was investigated over time-domain and frequency-domain feature vector for binary classification. The extracted and selected feature vectors were classified with various ML algorithms.

II. Materials & Methods

The block diagram representing the proposed approach is as in Fig. 1.

A. Dataset

In this study, processes were carried out using BCI Competition IV Dataset-IIa, which is a publicly available dataset in the literature [18]. This dataset includes EEG data from 9 subjects (4 female and 5 male). EEG signals were recorded over two sessions on different days, and four different MI tasks were performed: left hand, right hand, foot, and tongue. Each session includes six runs separated by short breaks, and each run also consists of 48 trials. These 48 trials were designed to be 12 trials per MI task. As a result, 288 trials were performed in one session for one subject. In recording of EEG signals, 22-channel EEG data were acquired using Ag/AgCl electrodes and sampled at 250 Hz. The signals are band-pass filtered between 0.5 Hz and 100 Hz and an additional 50 Hz notch filter is applied to suppress line noise.

B. Feature Extraction

The EEG segments where the MI tasks are performed are separated from the EEG signals and the time domain and frequency features are extracted to design a successful classification algorithm. The features used in the study are as follows based on time domain and frequency domain grouping.

- Time-domain features: Various features based on the amplitude and statistical changes of the EEG signals (Figure 1)
- Frequency-domain features: Energy, variance and entropy values of delta, theta, alpha, beta and gamma frequency bands

Time-domain feature extraction process use the 24 different feature including information about amplitude and statistical changes of the EEG signals [19]–[24]. Fig. 1 defined all time-domain features used in the study.

Frequency-domain feature extraction process use the frequency distribution embedded in the EEG signal. The EEG signal consists of five fundamental frequency bands based on the frequency range. These frequency bands are classified as delta (0.5-4Hz), theta (4-8Hz), alpha (8-13Hz), beta (13-30Hz) and gamma (>30) waves. In this study, each EEG channel was separated into frequency bands using FFT [19], and energy, variance and entropy values of delta, theta, alpha, beta and gamma bands were calculated.

Energy feature of each frequency band based on power spectrum calculated as

$$Energy = \sum_{i=1}^{M} y(i)^2 \tag{1}$$

Here, y corresponds to the Fourier transform of the signal and M corresponds to the maximum frequency. Variance of the each frequency band calculated as follows

$$Variance = \frac{1}{M-1} \cdot \sum_{i=1}^{M} (y_i - \overline{y})^2 \tag{2}$$

$\overline{y}$ indicates the average of the y signal.

The regularity of power spectrum of EEG signal measured and this is known as spectral entropy. The entropy values of each frequency band calculated as

$$Entropy = \frac{1}{log(M)} \cdot \sum_{i=1}^{M} P(y(i))log(P(y(i)) \tag{3}$$

$p(y(i))$ denotes the probability that the signal is in the given frequency range.

A data set consisting of 2592 samples from a total of 528 features was obtained by extracting 24 different time-domain features for each of the 22 EEG channels. In addition to the time domain features, 3 different frequency domain features were extracted for each frequency band from all 22 EEG channels, and a data set consisting of 2592 samples from a total of 330 features was obtained. Thus, the feature vector consists of a total of 858 features by combining 528 time domain and 330 frequency domain features.

C. Feature Selection based on Statistical Significance

The main purpose of feature selection is to define the features that will enable the highest discrimination between the groups of interest and minimize the classifier complexity [12]. In this paper, the statistical significance based feature selection used to find the best combination of time-domain and frequency-domain features which provide the best discrimination of the MI tasks for each sample. The statistical significance-based feature selection method used in this study was also applied in other BCI studies [12]. The independent t-test, that is mainly used to indicate significance of differences between measures of two different groups was used in the study because two MI tasks were tried to be classified. Thus, the effect of the independent t-test based feature selection method with time-domain and frequency domain EEG features on the performance of MI task classification was investigated. The p-values were calculated to observe statistical significances. The statistical significance level (α) define as 0.05 and the features that provide statistical evidence range were determined and applied as the input data for classifiers.

D. Classification

In the study, the performances of the feature sets, in which all features are used and statistically significant features are used, are tested with various classifier algorithms. Training and test data groups were separated using the 5-fold cross-validation method from the feature sets. 25 different classifier algorithms, which are widely used in the literature [25], were processed for the binary (right hand and left hand) MI task classification of feature sets. Each algorithm was evaluated 10 times to ensure the repeatability of the results. Average accuracy values of 10 evaluations calculated using true positives (TP), (true negatives) TN, false positives (FP) and false negatives (FN) values which are extracted from confusion matrix. Sensitivity and specificity performance metrics were also used for the algorithms for which the maximum accuracy value was calculated. The following equations were used in

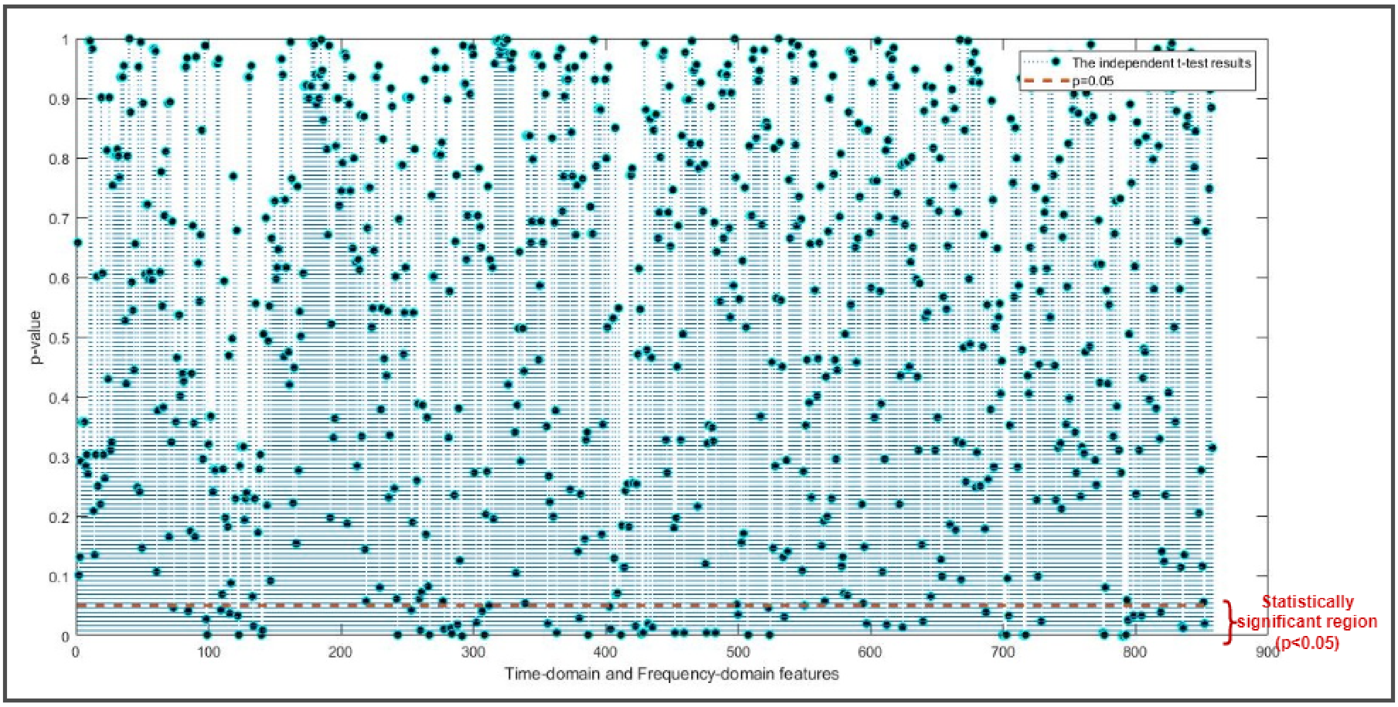

Figure 2: Statistical significance test results based on p-values of time-domain and frequency-domain features

the calculation of these metrics [11], [25]:

$$\text{Accuracy (\%)} = \frac{TP+TN}{TP+TN+FP+FN} \times 100 \quad (4)$$

$$\text{Sensitivity (\%)} = \frac{TP}{TP+FN} \times 100 \quad (5)$$

$$\text{Specificity (\%)} = \frac{TN}{TN+FP} \times 100 \quad (6)$$

The effect of independent t-test based feature selection on classifiers performance was examined using the repeated measures t-test. The paired-sample t-test compares the test results of all features applied and feature selected features. It was applied to examine whether there is a significant change between the first scenario and the second scenario based on p-value.

III. Experimental Results

In this study, a method based on the selection of statistically significant features and classification of this selected feature vector with different algorithms is proposed for MI task classification. The feature extraction and classification processes in the study were carried out in MATLAB application. The software package "IBM SPSS Statistics 25" which is probably the most common program in statistical analysis was used to apply independent t-test and find p-values that indicate statistical significance. In the first scenario, all time-domain and frequency-domain features were used, in the second scenario, statistically significant features were determined using independent t-test. The results of the independent t-test are given

	1st Scenario	2nd Scenario
Model	**Accuracy (%)**	
Fine Tree	53.60	53.50
Medium Tree	54.90	56.60
Coarse Tree	55.90	57.20
Linear Discriminant Analysis	54,90	61.70
Quadratic Discriminant	Failed	56.90
Logistic Regression	55.80	61.90
Gaussian Naive Bayes	50.20	56.20
Kernel Naive Bayes	52.50	55.90
Linear Support Vector Machine	60.42	61.18
Quadratic Support Vector Machine	**60.69**	58.54
Cubic Support Vector Machine	57.75	57.26
Fine Gaussian Support Vector Machine	49.80	49.20
Medium Gaussian Support Vector Machine	56.30	57.60
Coarse Gaussian Support Vector Machine	54.60	60.00
Fine K-Nearest Neighbours	49.70	52.30
Medium K-Nearest Neighbours	48.70	54.40
Coarse K-Nearest Neighbours	51.60	56.10
Cosine K-Nearest Neighbours	50.10	54.40
Cubic K-Nearest Neighbours	49.60	53.50
Weighted K-Nearest Neighbours	49.00	54.30
Ensemble Boosted Trees	59.00	59.20
Ensemble Bagged Trees	53.83	57.10
Ensemble Subspace Discriminant	59.87	**62.52**
Ensemble Subspace K-Nearest Neighbours	53.48	53.35
Ensemble RUSBoosted Trees	54.36	56.60

Table I: Binary classification performance of the time-domain and frequency-domain features and effectiveness of the independent t-test based feature selection.

Study	Dataset	Number of Channels	Feature extraction	Feature selection	Classes	Classifier	ACC (%)	SEN (%)	SPE (%)
[10]	BCI Competition IV Dataset-IIa	2 EEG (C3 and C4)	WPD	DFFS	Left hand, Right hand	RF	68.32	N/A	N/A
[15]	BCI Competition IV Dataset-IIa	22 EEG	MEMD and Sample covariance matrix	N/A	Left hand, Right hand	Riemannian geometry	79.93	N/A	N/A
[16]	BCI Competition IV Dataset-IIa	22 EEG	WT and CSP	N/A	Left hand, Right hand, Feet, Tongue	HSVM	64.40	N/A	N/A
[17]	MISCP	21 EEG	Multi-class CSP	N/A	Five fingers	SVM	40.60	N/A	N/A
[26]	BCI Competition IV Dataset-IIa	22 EEG	Time-domain features	ANOVA	Left hand, Right hand, Feet, Tongue	Linear Discriminant Analysis	44.00	44.30	44.31
This study	BCI Competition IV Dataset-IIa	22 EEG	Time-domain and Frequency-domain features	Independent t-test	Left hand, Right hand	ESD	62.52	64.51	60.65

Table II: Comparison of motor imaginary task classification studies with the results of the proposed study. ACC, SEN, and SPE are the accuracy, sensitivity, and specificity, respectively.

in Fig. 2 for all features. The red line indicates p=0.05 value, p-values below that line indicate significant and discriminative features. According to the results of the statistical analysis based on the independent t-test, 71 out of 858 features yielded a significant p-value. The extracted and selected feature sets were tested 10 times using 25 different classifier algorithms and 5-fold cross-validation. Average accuracy values of all repeated tests were evaluated and are given in Table I. In the first scenario, the highest 60.69% average accuracy value was achieved using Quadratic Support Vector Machine (SVM) algorithm and in the second scenario, the highest 62.52% average accuracy value was achieved using the Ensemble Subspace Discriminant (ESD) algorithm developed under the Ensemble Learning classification.

IV. Discussion and Conclusion

The aim of this study is to investigate the effects of statistically significant time-domain and frequency-domain features on binary classification of MI tasks. According to the Table I, when the scenario in which the features selected as statistically significant by independent t-test is used is compared with the scenario where all the features are used, it is seen that the performance increased in 20 classifiers, and the performance decreased in 5 classifiers. According to the classifier performances, the independent t-test based feature selection process generally improves the performance of the the classifier but it is difficult to reach a general decision. Therefore, the repeated measures t-test was used to check whether significant difference between scenario 1 and scenario 2. The paired-sample t-test applied to compare the results of scenarios. The p-value was calculated to determine statistical significance. The paired-sample t-test found the p value of 0.04. As p value less than 0.05 indicates that there is a significant difference between the two groups compared. As a result, the independent t-test is effective feature selection method that improves the classifier performance and using of this method generates a significant difference in classification.

When the experimental results obtained in this study are compared with other studies in the literature that classify binary and multi-class MI tasks (Table II), it is seen that the highest classifier performance obtained in this study is lower than other studies. However these studies examined in detail, it was observed that either a more complex feature selection algorithm was used or EEG channel selection was applied simultaneously in other studies [10]. In multi-task classifications, the discernibility of motor imaginary tasks decreases [16], [17], [26]. In the Table II, the information about which EEG channels, the feature extraction methods, the feature selection algorithms and classifier algorithms were used in the related studies are also given in detail. In further studies, it is aimed to improve the classification performance by using channel selection, different feature extraction methods and effective feature selection algorithms such as GA.

References

[1] Musallam YK, AlFassam NI, Muhammad G, Amin SU, Alsulaiman M, Abdul W, Altaheri H, Bencherif MA, Algabri M. Electroencephalography-based motor imagery classification using temporal convolutional network fusion. Biomedical Signal Processing and Control 2021; 69: 102826.

[2] Panicker RC, Puthusserypady S, Sun Y. An asynchronous p300 bci with ssvep-based control state detection. IEEE Trans Biomed Eng 2011; 58(6): 1781–1788.

[3] Degirmenci M, Sayilgan E, Isler Y. Evaluation of wigner-ville distribution features to estimate steady-state visual evoked potentials' stimulation frequency. Journal of Intelligent Systems with Applications 2021; 4(2): 133-136.

[4] Pawar D, Dhage S. Feature Extraction Methods for Electroencephalography based Brain-Computer Interface: A Review. International Journal of Computer Science (IAENG) 2020; 47(3).

[5] Djamal EC, Abdullah MY, Renaldi F. Brain computer interface game controlling using fast fourier transform and learning vector quantization. Journal of Telecommunication, Electronic and Computer Engineering (JTEC) 2017; 9(2-5): 71-74.

[6] Xu B, Zhang L, Song A, Wu C, Li W, Zhang D, Xu G, Li H, Zeng H. Wavelet transform time-frequency image and convolutional network-based motor imagery EEG classification. IEEE Access 2018; 7: 6084-6093.

[7] Luo J, Gao X, Zhu X, Wang B, Lu N, Wang J. Motor imagery EEG classification based on ensemble support vector learning. Computer Methods and Programs in Biomedicine 2020; 193: 105464.

[8] Wang J, Feng Z, Lu N, Sun L, Luo J. An information fusion scheme based common spatial pattern method for classification of motor imagery tasks. Biomedical Signal Processing and Control 2018; 46: 10-17.

[9] Ang KK, Chin ZY, Wang C, Guan C, Zhang H. Filter bank common spatial pattern algorithm on BCI competition IV datasets 2a and 2b. Frontiers in Neuroscience 2012; 6: 39.

[10] Luo J, Feng Z, Zhang J, Lu N. Dynamic frequency feature selection based approach for classification of motor imageries. Computers in Biology and Medicine 2016; 75: 45-53.

[11] Isler Y. A Detailed Analysis of the Effects of Various Combinations of Heart Rate Variability Indices in Congestive Heart Failure, PhD Thesis, Dokuz Eylul University, 2009.

[12] Narin A, Isler Y, Ozer M. Investigating the performance improvement of HRV Indices in CHF using feature selection methods based on backward elimination and statistical significance. Computers in Biology and Medicine 2014; 45: 72-79.

[13] Mousa MA, El-Khoribi RA, Shoman ME. A novel brain computer interface based on principle component analysis. Procedia Computer Science 2016; 82: 49-56.

[14] Selek MB, Yesilkaya B, Egeli SS, Isler Y. The effect of principal component analysis in the diagnosis of congestive heart failure via heart rate variability analysis. Proceedings of the IMechE Part H: Journal of Engineering in Medicine 2021; 235(12): 1479-1488.

[15] Gaur P, Pachori RB, Wang H, Prasad G. A multi-class EEG-based BCI classification using multivariate empirical mode decomposition based filtering and Riemannian geometry. Expert Systems with Applications 2018; 95: 201-211.

[16] Dong E, Li C, Li L, Du S, Belkacem AN, Chen C. Classification of multi-class motor imagery with a novel hierarchical SVM algorithm for brain–computer interfaces. Medical & Biological Engineering & Computing 2017; 55(10): 1809-1818.

[17] Kato M, Kanoga S, Hoshino T, Fukami T. Motor imagery classification of finger motions using multiclass CSP. In 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), July 20-24, 2020, Montreal, QC, Canada, pp. 2991-2994.

[18] Brunner C, Leeb R, Muller-Putz G, Schlogl A, Pfurtscheller G. BCI Competition 2008–Graz data set A. Institute for Knowledge Discovery (Laboratory of Brain-Computer Interfaces), Graz University of Technology 2008; 16: 1-6.

[19] Sayilgan E, Yuce YK, Isler Y. Frequency recognition from temporal and frequency depth of the brain-computer interface based on steady-state visual evoked potentials. Journal of Intelligent Systems with Applications 2021; 4(1): 68-73.

[20] Sayilgan E, Yuce YK, Isler Y. Evaluation of mother wavelets on steady-state visually-evoked potentials for triple-command brain-computer interfaces. Turkish Journal of Electrical Engineering & Computer Sciences 2021; 29(5): 2263-2279.

[21] Sayilgan E, Yuce YK, Isler Y. Investigating the effect of flickering frequency pair and mother wavelet selection in steady-state visually-evoked potentials on two-command brain-computer interfaces, Innovation and Research in BioMedical Engineering 2022; IN PRESS.

[22] Sayilgan E, Yuce YK, Isler Y. Evaluation of wavelet features selected via statistical evidence from steady-state visually-evoked potentials to predict the stimulating frequency. Journal of the Faculty of Engineering and Architecture of Gazi University 2021; 36(2): 593-605.

[23] Sayilgan E, Yuce YK, Isler Y. Determining gaze information from steady-state visually-evoked potentials. Karaelmas Science and Engineering Journal 2020; 10(2): 151-157.

[24] Sayilgan E, Yuce YK, Isler Y. Estimation of three distinct commands using Fourier transform of steady-state visual-evoked potentials. Duzce Universitesi Bilim ve Teknoloji Dergisi 2020; 8(4): 2337-2343.

[25] Hart PE, Stork DG, Duda RO. Pattern Classification, A Wiley-Interscience Publication, 2001.

[26] Degirmenci M, Yuce YK, Isler Y. Motor imaginary task classification using statistically significant time-domain EEG features. In 2022 30th Signal Processing and Communications Applications Conference (SIU), May 16-18, 2022, Safranbolu, Turkey, ACCEPTED.