

Article

Classification of Motor Imagery Using a Combination of User-Specific Band and Subject-Specific Band for Brain-Computer Interface

Vacius Jusas * and Sam Gilvine Samuvel

Department of Software Engineering, Kaunas University of Technology, LT-51390 Kaunas, Lithuania; Sam.samuvel@ktu.edu

* Correspondence: Vacius.jusas@ktu.lt

Received: 14 October 2019; Accepted: 16 November 2019; Published: 20 November 2019

Abstract: The essential task of a Brain-Computer Interface (BCI) is to extract the motor imagery features from Electro-Encephalogram (EEG) signals for classifying the thought process. It is necessary to analyse these obtained signals in both the time domain and frequency domains. It is observed that the combination of multiple algorithms increases the performance of the feature extraction process. This paper identifies combinations that have not been attempted previously and improves the accuracy of the overall process, although other authors implemented different combinations of the techniques. The focus is given more on the feature extraction process and frequency bands, which are user-specific and subject-specific frequency bands. In both time and frequency domains, after analysing EEG signals with the time domain parameter, we select the frequency band and the timing while using the Fisher ratio of the time domain parameter (TDP). We used Fisher discriminant analysis (FDA)-type F-score to simultaneously select the frequency band and time segment for multi-class classification. We extracted subject-specific TDP features from the training trials to train the classifier when optimal time-frequency areas were selected for each subject. In this paper, various methods are explored for obtaining the features, which are Time Domain Parameters (TDP), Fast Fourier Transform (FFT), Principal Component Analysis (PCA), R^2 , Fast Correlation Based Filter (FCBF), Empirical Mode Decomposition (EMD), and Intrinsic time-scale decomposition (ITD). After the extraction process, PCA is used for dimensionality reduction. An efficient result was obtained with the combination of TDP, FFT, and PCA. We used the multi-class Fisher's linear discriminant analysis (LDA) as the classifier, which was in line with the FDA-type F-score. It is observed that the combination of feature extraction techniques to the frequency bands that were selected by the Fisher ratio and FDA type F-score along with Fisher's LDA classifier had higher accuracy than the results obtained other researches. A kappa coefficient accuracy of 0.64 is obtained for the proposed technique. Our method leads to better classification performance when compared to state-of-the-art methods. The novelty of the approach is based on the combination of frequency bands and two feature extraction methods.

Keywords: Motor Imagery; feature extraction; user specific band; subject specific band

1. Introduction

The possibility of controlling other devices while using brain functions has become feasible. The brain signals can be intercepted through the process of Brain Computer Interface (BCI). This is possible by recording the various signals from the brain and analysing them to identify the type of the signal based on the time and frequency of the signal [1]. The signals can be internally and externally collected from the brain. It is more effective in collecting these signals from internal measurement, since the measured values are more accurate. The externally measured signals are of

comparatively lower quality when compared to the internally measured values [2]. However, internal measurements are a potentially dangerous technique, since the measuring devices have to be surgically placed inside the body, leading to higher chances of life threat and, hence, it is also known as an invasive method. The externally measured non-invasive method does not require any form of surgical procedure and is more preferred in spite of its comparatively lower quality. Some of these non-techniques are functional Magnetic Resonance Imaging (fMRI), Computed Tomography (CT), Magneto-Encephalo-Graphy (MEG), and Electro-Encephalo-Graphy (EEG).

When these techniques are compared, EEG is a better technique, since it is easier to use, lower investment cost, and has a comparatively superior temporal resolution [3,4]. Hence, it is the most popular BCI measurement technique. A measuring device is placed on the scalp, where it detects the electrical neural signals and identifies the type of patterns, like the state of the subject. It identifies whether the subject is calm, hypnotised, or angry.

Identifying the state of the subject from the signal is challenging and, hence, the classification must be done appropriately by initially training the algorithm efficiently. The BCI has lots of advantages and it has the capability of using it in a variety of applications. This interface can be used by the physically challenged for controlling the directions of the wheelchair. Another application is where the computer input functions, like the mouse pointer or the cursor of the computer, can be controlled.

Motor Imagery is the process where these simulations are performed while using brain. The appropriate frequency bands that present these features are Beta band and Mu band in the EEG systems. These frequency bands that include the important features vary according to the individual and also depend upon the event dependent property. Hence, it is necessary to analyse the individual EEG signals and with respect to the time domain and frequency domain, so that the characteristics can be learnt. There are lots of feature extraction techniques and classification techniques available that have been performed by various researchers. However, the accuracy of the approaches is not up to the mark and, hence, the novel techniques must be identified for increasing the accuracy. This paper has compared different feature extraction techniques and classification techniques by combining them in various combinations in order to identify the most efficient combination. Our contribution of this work is combining user specific band and subject specific band and two feature extraction methods together.

In this research, we studied different time and frequency intervals for each subject and determined the best subject specific time and intervals for the classification of four-class motor imagery tasks. Different time intervals were examined with various feature extraction methods and classifiers to find out the best time intervals each subject who performed motor imagery. User specific and subject specific bands are both effective in performance improvement. This paper aims to show using combination of user specific and subject specific bands would have better performance than just using user specific and subject specific bands alone. Our proposed method yielded better classification performances than the state-of-the-art methods.

This section introduces the motor imagery and EEG for controlling the Brain Control Interface (BCI). The next section discusses the literature review, where the related studies are reviewed and compared in order to identify the gap and the most suitable combination of the techniques. This is followed by the methodology, where the selected dataset and the various combinations of feature extractions are showed. This is followed by the obtained results and they are then interpreted. Finally, the paper is concluded, and the optimum combination of techniques is identified.

2. Related Work

A hybrid method employing common spatial pattern filter has been used by Bais et al. [5] for extracting the features. Different optimisation techniques, like Particle Swarm Optimisation (PSO), Ant Colony Optimisation (ACO), Artificial Bee Colony (ABC), and Simulated Annealing. These methods have been used along with support vector machine (SVM) classifier; however, the performance is slower when compared to the conventional techniques.

Dimensionality reduction techniques have been investigated by García-Laencina et al. [6] by using the Hjorth parameters and adaptive autoregressive coefficients. The feature selection techniques that were used were sequential forward selection and sequential backward selection, while the classification techniques used were Principal Component Analysis (PCA), Local Fisher Discriminant Analysis (LFDA), and Locality Preserving Projections. The performance of the classification process has been increased using these hybrid techniques. However, the number of channels and subjects used are very less and, hence, more detailed analysis is necessary.

Wavelet transform was used by Gupta et al. [7] for extracting the features. Six types of filter methods, which are Euclidean distance (ED), Bhattacharyya distance measure (BD), Kullback–Leibler distance (KD), ratio of scatter matrices (SR), linear regression (LR), and maximum relevance minimum redundancy (mRMR), have been used for decreasing the size of feature vectors. Each of the six feature selection techniques has increased the performance of the classification. From these techniques, the combination of wavelet transform and linear regression is seen to have the best performance.

A feature reduction algorithm has been proposed by Han et al. [8] for the EEG framework. Pre-processing is performed while using autoregressive coefficients. The three-dimensional EEG signals are converted to two-dimensional matrix by compressing the feature vectors. Different techniques have been used for the ranking process, such as RFS, SSLSR, and RUFFS. However, only two motor imagery tasks have been performed in this work.

Jusas and Samuvel [9] has been classified the Motor Imagery while using combinations of feature extraction and dimensionality reduction approaches. FFT, TDP, band power, and channel variance are the methods used by the researchers. The obtained values are combined together in pairs. Different feature reduction techniques have been analysed, such as PCA, sequential selections, LPP, and LFDA. It has been observed that the combined techniques have higher accuracy than the individual techniques. Hence, it compares the various combination of techniques for improved accuracy and recommends the FFT, CV, and PCA methods along with the LS-SVM classification. Hence, this work will combine the different techniques for improving the accuracy.

Mahmoudi and Shamsi [10] proposed a method for finding the subject specific time intervals for the classification of four-class motor imagery tasks by using mutual information (MI) between the BCI input and output. The signal-to-noise ratio was utilized to compute the MI values while the MI values were utilized as feature selection criteria to select the discriminative features. The time segments and the better discriminative features were found by using training data and used to estimate the evaluation data. The filter bank common spatial pattern (FBCSP) algorithm has been divided in to four progressive stages, such as filter-bank, the CSP algorithm, feature selection, and classification. However, the noise and artifacts of the EEG signals have been unnoticed in the experiment.

Ren et al. proposed a feature extraction technique that has combined the feature extraction and feature selection method [11]. Four different feature extraction methods were used to reduce the number of features from a total of 83 features. Three different feature selection techniques were used and then compared. The Fisher score has been seen to have the most accurate results. The other techniques were seen to have a comparatively lesser accuracy.

Rodríguez-Bermúdez et al. proposed a wrapper based methodology [12] for selecting the features. The features used are power spectral density features along with the AAR co-efficient and Hjorth parameters. These averages of these features are calculated and then compiled into a single vector. These features are then selected and subjected to different regression techniques and it has been seen that the Least Angle Regression (LARS) algorithm works better than Wilcoxon rank test.

Wang et al. [13] presented a statistical model to select the optimal feature subset based on the Kullback-Leibler divergence measure and automatically select the optimal subject-specific time segment. The autoregressive model and log variance are employed on the Common Spatial Patterns (CSP) for the feature extraction. These extracted factors are spatial and temporal correlated power spectral features. In the experiment, they only performed binary classification. We used a four-class motor imagery tasks in this research.

Yuan et al. [14] proposed FDA techniques for reducing the linear dimensionalities, where the FDA and linear discriminant analysis (LDA) were combined. A graph that is data adaptive is created with the L1 or L2 norm constraints and then merged with the LDA approach for a better analytic solution. The experimental results are implemented on different datasets to demonstrate the effectiveness. It is seen to be effective in lower dimensions and small training datasets.

Yu et al. [15] used spatial filter techniques along with PCA. It has been seen that common average reference filter performs better than other well-known spatial filter techniques. However, the feature reduction using PCA did not improve the accuracy, but the performance of the classification was maintained.

Zhang et al. combined an autoregressive model and sample entropy model [16] for extracting the features. The coefficients of these models have been used along with SVM and Radial Basis Function (RBF) for the classification. The obtained accuracy was higher than the individual autoregressive models; however, the obtained accuracies were lower than existing combined techniques.

Uktveris and Jusas [17] considered a deep learning approach based on convolutional neural network (CNN). CNN and their application to four-class motor-imagery based problem were analysed in the research. The experimental results are similar to more complex state-of-the-art EEG analysis techniques.

Dai et al. [18] proposed an approach that combines CNN and variational autoencoder (VAE) networks. Deep learning approaches were used in the experiment. The deep learning approach takes extremely large amount of data to perform better than other methods [17]. Hence, it is difficult to obtain better results in the deep learning approach. There will be lesser chance to improve the classification accuracy.

3. Methods

In this paper, two feature extraction processes are performed to increase the number of relevant features. When lots of features are used, it leads to errors and confusion, thereby reducing the efficiency. Therefore, PCA method is used for the feature reduction. Hence, a CSP is used for the feature decomposition. The following feature extraction methods, such as time domain parameters, empirical mode decomposition, fast fourier transform, fast correlation-based filter, intrinsic time-scale decomposition, and squared pearson's correlation, have been combined in several ways.

The Fisher's LDA and Least Square Support Vector machine (LS-SVM) are used for the classification purpose, as shown in Figure 1. The frequency bands are selected based on the users and subjects. The Fisher's ratio of the time domain parameters is calculated to identify the dominant frequency band and timing in the signals. This is performed by applying the by-band pass filter.

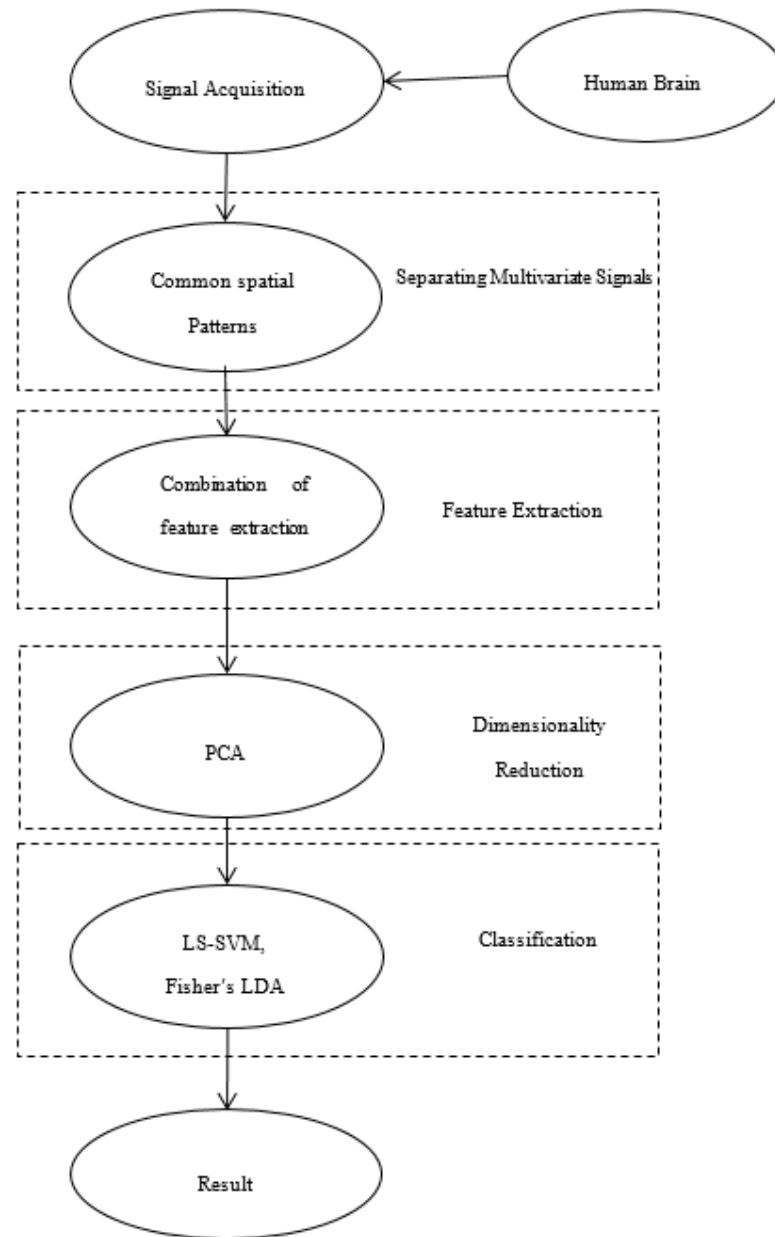


Figure 1. A block diagram of proposed approach.

A relatively new and perspective approach to motor imagery was found in combination of feature extraction methods. A combination of user specific band and subject specific band is a novel method based technique that has not been used with EEG. The combination of user specific band and subject specific band could be the new perspective way to present a solution since EEG motor imagery task lacks accurate solutions.

3.1. Common Spatial Patterns

CSP was initially used for classifying multiple channels EEG by Ramoser [19]. It was mainly used for linear transformation for projecting multiple channel EEG data into low dimensional spatial space by using a projection matrix, where every row contains weights for channels. The transformation can increase the variance of dual class signal matrices. This method uses and diagonalizes the co-variance matrices of both classes [20]. Let X_1 of size (n, t_1) and X_2 of size (n, t_2)

be two multivariate signal windows, where n is the number of signals and t_1 and t_2 are the number of samples respectively. The CSP determines the W^T component, so that the ratio of variance is maximized between the two windows. This can be expressed as follows:

$$w = \arg \max_w \frac{\|wX_1\|^2}{\|wX_2\|^2} \tag{1}$$

3.2. Fisher Ratio

A feature is computed by utilizing the time domain parameter at each k -th window. “Along this way, whole features are obtained in all of the windowed durations of training EEG signals, and the features of each class are then ensemble averaged. Additionally, these processes are repeated, changing a frequency band to others” [21]. The following way is to select the frequency bands. Within the frequency range of 5–30 Hz, the n frequency points are defined, and a frequency band is composed of two points among n points. Subsequently, the number of bands is that the number of 2 combinations from frequency points nC_2 . The Fisher ratio, where j is the filter index, k is the window index, and l is the index of the time domain parameter, is calculated from the averaged features of two classes in order to select the important timing and frequency band, as follows:

$$F(j, k, l) = \frac{|m_1(j, k, l) - m_2(j, k, l)|}{\sigma_1(j, k, l)^2 + \sigma_2(j, k, l)^2} \tag{2}$$

where, $m_i(j, k, l)$ ($i = 1, 2$) denotes the average and $\sigma_i(j, k, l)^2$ stands for the variance of the time domain parameter l of each i at k -th window and j -th filter.

3.3. FDA- Type F-Score

FDA-type F-score is a simplified measure that is based on Fisher discriminant analysis (FDA) for assessing the discriminative power of a group of features (a feature vector) [22].

$$F = \frac{\|\vec{\mu}^L - \vec{\mu}^R\|_2^2}{tr(\Sigma^L) + tr(\Sigma^R)} \tag{3}$$

In above equation, where, $\vec{\mu}$ denotes the mean of the feature vector, Σ denotes the covariance matrix of the feature vector, and tr denotes the trace of a matrix. “Thus, FDA-type F-score depend on the Euclidean distance between class centers to evaluate the difference between classes and utilizes the trace of the covariance matrix to estimate the variance within one class”. FDA-type F-score, as a simplified criterion, avoids estimating a projection direction in multi-dimensional FDA, and it has been efficiently used in two class BCI and motor recognition studies for channel and feature selection.

3.4. Feature Extraction Techniques

There are lots of feature extraction approaches that can be used for extracting the EEG signals. The extraction techniques used in this paper are TDP, FFT, and PCA.

3.4.1. Time Domain Parameters

Time Domain parameters (TDP) is a technique that is performed by calculating the time-varying power of first k derivatives in the following equation:

$$p_i(t) = \left(\frac{d^i x(t)}{d_t^i} \right), \dots \dots i = 0, 1, 2, \dots, k \tag{4}$$

The values that are obtained can be smoothened by utilising an exponential moving average window filter. Even though the features of the TDP are defined in the time domain, they can also be inferred as frequency domain filters. Other spectral approaches that are available are Fourier transform, wavelet analysis, and autoregressive spectrums, which can define the rest of the spectral density function. However, this technique has the limitation that classification does not occur when

there are too many parameters. Therefore, the training requires lots of data that lead to the possibility of overfitting [23].

3.4.2. Empirical Mode Decomposition

Empirical Mode Decomposition (EMD) is an adaptive signal analysis technique for analysing the signal with a wide range of applications. It decomposes the signal into unique and different frequency components, known as Intrinsic Mode Functions (IMF). If this decomposition is not possible, other mode functions will then contain similar frequencies as overlapping components [24]. There are certain criteria to be applied as an IMF. The total zero crossings in the data must not be more than one. It should preferably be the same. This means that the values must remain either positive or negative. Its polarity must not change. The average value of envelope from maxima and minima must be zero at all times. The EMD has the ability to convert any signal to IMF. It is a shifting process performed to decompose the signal into narrow band signals. The criteria are satisfied by the expression below:

$$h_{1(i-1)} - m_{1i} = h_{1i} \quad (5)$$

3.4.3. Fast Fourier Transform

Fast Fourier Transform (FFT) analyses certain signals and samples for certain space and time. It splits them into smaller frequency components. It processes the discrete Fourier transform for few data samples. This data that should be transformed is partitioned into smaller frames [9]. Each individual frame is transformed and the obtained result is added to the matrix. Short-Time Fourier Transform (STFT) is a type of Fourier transform that can be represented by following equation:

$$STFT = X(m, w) = \sum_{-\infty}^{\infty} x[n] w[n - m] e^{-jwn} \quad (6)$$

Here, w is continuous and m is discrete. This is performed using FFT; hence, both of these variables are quantized and discrete. The FFT is very consistent and robust in obtaining the most optimum features.

3.4.4. Fast Correlation Based Filter

FCBF is a multivariate type of feature selection technique that uses Symmetrical Uncertainty (SU) for calculating the dependencies of the features and identifying the best subset while using backward selection technique along with sequential search strategy. It contains internal conditions, where the process stops when the necessary criteria is satisfied. It is based on correlation that generally runs faster than the other subset selecting techniques. Entropy and conditional entropy values are used for calculating the feature dependencies [25]. The entropy is calculated by the following expression:

$$H(X) = \sum p(Xi) + \log_2(p(Xi)) \quad (7)$$

where, x is a random variable and $p(x)$ is its probability.

3.4.5. Intrinsic Time-Scale Decomposition

Intrinsic time-scale decomposition (ITD) is a type of signal processing technique that has been recently developed. This technique can split a complicated signal into multiple smaller Proper Rotation Components (PRC) on the basis of local time scale of the signal characteristics. The signals are determined at the local extremum point by using a linear transformation technique; hence, more local data can be used for the process [26]. The Intrinsic time-scale decomposition can be computed while using the formula:

$$X_t = LX_t + (1 - L)X_t = L_t + H_t \quad (8)$$

3.4.6. Squared Pearson's Correlation

It is the proportion of the variance when the dependent variable can be obtained from the independent variable. It is denoted by R^2 and it is the square of the correlation between the actual and predicted outcomes. It is used in statistical analysis for regression and it can be used for various types of analysis. This method independently estimates the discriminative power of every feature by calculating the square value of the Pearson's correlation coefficient between the values of the j th feature and the class vectors [27].

$$r_j = \frac{\sum_{i=1}^N (x_{ji} - \bar{x}_j)(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_{ji} - \bar{x}_j)^2 \sum_{i=1}^N (y_i - \bar{y})^2}} \quad (9)$$

In above equation, x_{ji} denotes the i th sample of j th feature, y_i denotes the class label associated with the i th sample, and the bar notation denotes the average value across all samples.

3.5. Feature Reduction

Principal Component Analysis

PCA is another technique that can be used for extracting the features through the filtering technique. It uses an orthogonal transformation for converting some observations of correlated variables to a group of uncorrelated variables [28]. It is an unsupervised method that calculates the linear mapping for accomplishing low size representation of original data, where there is a high amount of variance. The covariance of two variables X and Y is obtained while using the following equation:

$$Cov(X, Y) = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{x})(Y_i - \bar{y}) \quad (10)$$

The following processes take place in PCA:

- The co-variance matrix of the data points is obtained.
- The individual eigen values are calculated and then sorted in decreasing order.
- The first k -eigen vector is selected and this will have k dimensions.
- The original set of dimensions is modified. i.e., the dimensions are transformed into k dimensions.

3.6. Classification

3.6.1. Fisher's LDA

Fisher's Linear Discriminant Analysis, or just LDA, is a technique used for recognizing the patterns in statistics and machine learning for identifying various features that segregates different classes of objects and events. This combination of features can be used as a linear classifier for reducing the dimensionalities before the final classification.

The transformation in this algorithm is based on increasing the ratio of the variances of between the classes to within class with the aim of reducing the differences in the data within the class and increasing the differences between the classes [14]. This technique works well for multiple class problems. When there are X number of classes, the technique uses the $(X-1)$ projections by using the projection vectors θ_i , arranged by column matrix, where expressed in following equation:

$$y_i = \theta_i^T X \quad (11)$$

3.6.2. LS-SVM

The Least-squares Support Vector Machine (LS-SVM) is a least square version of SVMs that can analyse the data and identify recognisable patterns for classifying the data [29]. An SVM uses data

points as input and gives the output in form of a hyper-plane. A decision boundary is used for the classification between the different classes, which will be classified. Instead of quadratic programming, like conventional SVMs, the LS-SVM technique uses linear equations for solving them and they are a type of kernel based learning method [30].

4. Experimental Studies

4.1. Data

2a dataset from BCI competition IV is used for the analysis. The available data are collected from nine different people in two seasons and has been recorded on different days. The participants were given four different motor imagery tasks, like actual movement and imagination of different parts of the body, like hands, feet, and tongue. Totally, there were 288 trials performed with 72 trials in each class randomly. The dataset contains 22 EEG signals that are recorded in a monopolar manner. The signals were sampled at the frequency of 250 Hz and a band pass filter is applied to remove the lower order frequency. Figure 2 shows the structure of a single trial.

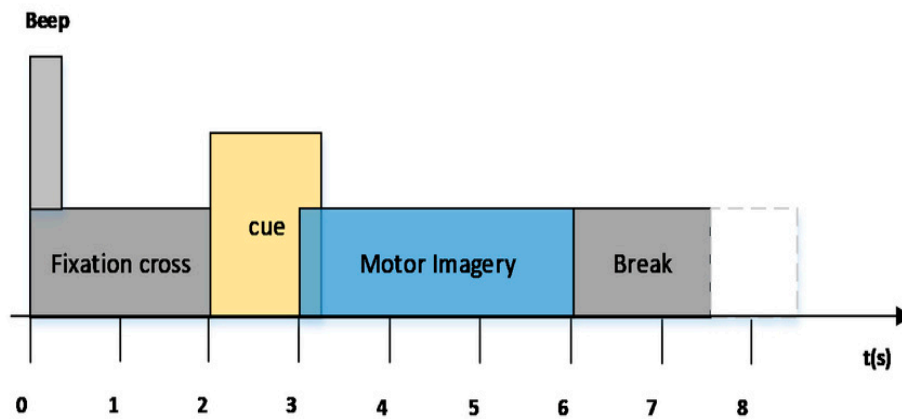


Figure 2. The time course of a single trial in the Brain Control Interface (BCI) competition IV dataset 2a.

4.2. Results

The classification results are evaluated and compared by utilizing the kappa coefficient, which takes the value 0 for a random classifier and 1 for a perfect classifier that consistently correctly classifies. The estimation of kappa coefficient is computed utilizing the equation below:

$$K = \frac{P_0 - P_e}{1 - P_e} \quad (12)$$

where P_0 denotes the classification accuracy and P_e denotes the hypothetical accuracy of a random classifier on the same data.

$$A = \frac{(P_0 - 0.25)}{1 - 0.25} \quad (13)$$

In above equation, we consider the value for $P_e = 0.25$. The final proportion of execution of a given algorithm is the maximum value of the kappa value from the computed time-course.

The Fisher ratio is performed for obtaining the user specific band. After analysing the EEG signals and comparing it with the time domain parameters, the most dominant frequency band and timing is identified by using the Fisher ratio of the time domain parameter. The FDA type F-score has been used for the subject specific band. Fisher's discriminant analysis is performed to identify both the dominant frequency band and timing for multiple class classification. It estimates the time and frequency areas for extracting the features of TDP with respect to subject. While the user

specific and subject specific band are individually applied, they are also applied together as the proposed system. Different frequency intervals and time intervals are observed to identify the most optimum intervals. To identify the best time interval, they are examined with the feature extraction techniques and feature reduction while using PCA. After this, Fisher's LDA was used as a classifier.

Figure 3 shows an original EEG signal. The electroencephalogram (EEG) is the recording of the electrical movement of the brain from the scalp. The recorded waveforms indicate the cortical electrical movement. EEG activity is quite small and they are measured in terms of microvolts.

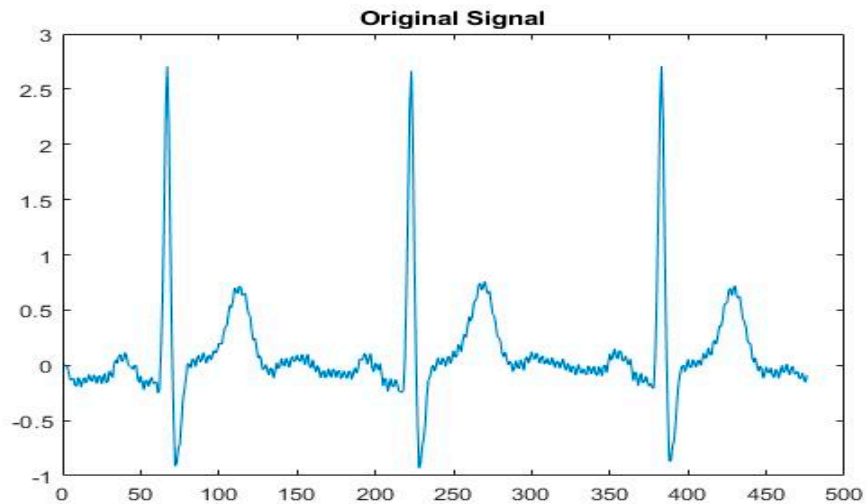


Figure 3. Original Electro-Encephalo-Graphy (EEG) signal.

Figure 4 shows a band-pass filter is used to display frequencies that are either too low or too high, making it easy to pass frequencies within a certain range. Band-pass filters can be created by stacking a low-pass filter at the end of a high-pass filter.

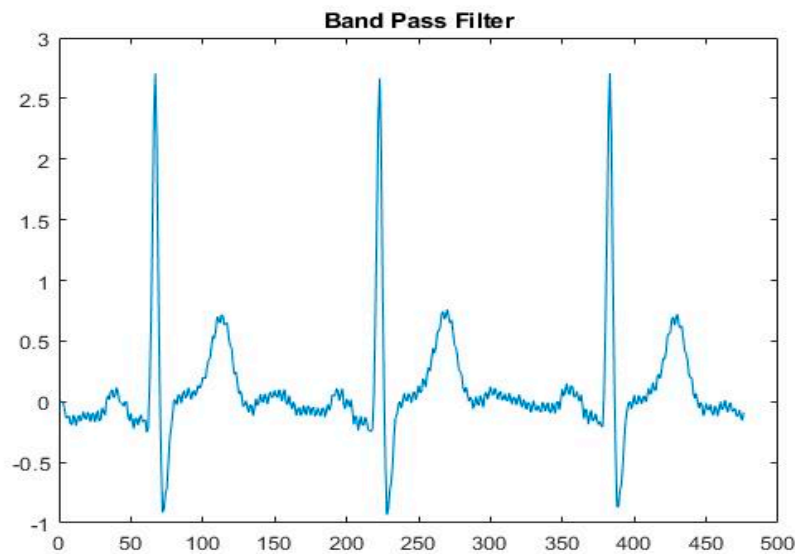


Figure 4. Band pass filter.

The Daubechies wavelets are a family of orthogonal wavelets and identify discrete wavelet transform. The coefficients of a one-dimensional signal are reconstructed and they are shown in Figure 5. The brain has five distinctive categories of brain waves; Gamma, Theta, Delta, Alpha, and Beta brain waves.

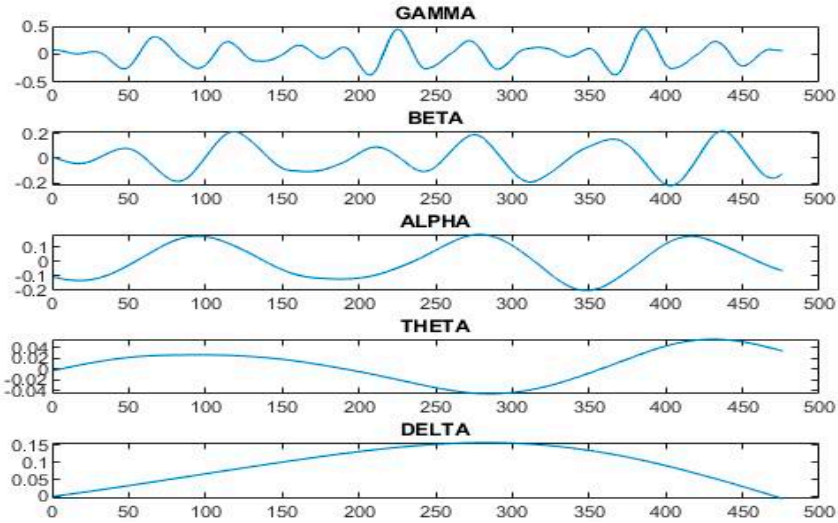


Figure 5. EEG brain waves.

Figure 6 shows the frequency of brain waves. Gamma wave occurs maximum frequency at 12.00 Hz. Beta wave occurs maximum frequency at 6.00 Hz. Alpha wave occurs maximum frequency at 3.00 Hz. Theta wave occurs maximum frequency at 1.00 Hz. Delta wave occurs Maximum frequency at 1.00 Hz.

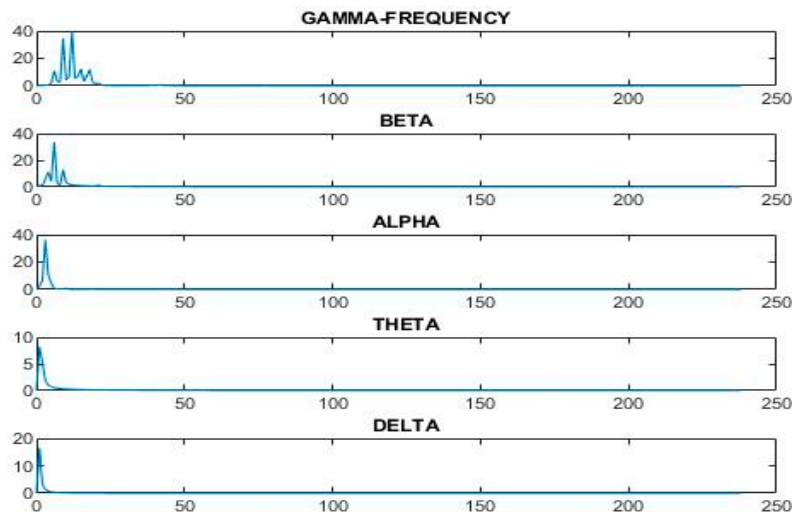


Figure 6. Frequency in EEG brain waves.

Band pass filter is utilized for EEG signal denoising. Figure 7 shows the denoised EEG signal.

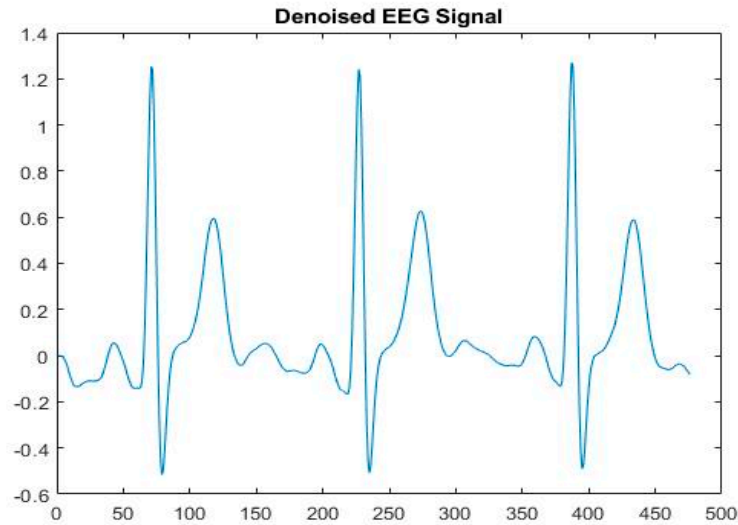


Figure 7. Denoised EEG signal.

This work performs a ten-fold cross validation. The TDP is utilised for identifying the optimum frequency band and timing during the training process. The band is subject to by-pass filtering to remove undesirable noises. Table 1 gives the range of the optimum frequency band and timing.

Table 1. Frequency band and their respective time.

Subject	Left Hand		Right Hand		Feet		Tongue	
	Frequency Band (HZ)	Time (sec)	Frequency Band (HZ)	Time (sec)	Frequency Band (HZ)	Time (sec)	Frequency Band (HZ)	Time (sec)
1	5–14	7.5–8.5	12–24	6.6–7.5	5–14	6–7	12–24	6.5–7.5
2	5–14	7.5–8.5	12–24	7.5–8.5	12–24	6–7	5–14	5.5–6.5
3	8–30	5.5–6.5	8–20	5.5–6.5	8–30	5.5–6.5	12–24	5–6
4	8–30	6.5–7.5	8–20	6–7	12–24	5–6	8–30	5–6
5	8–20	6–7	12–24	6–7	8–20	5–6	5–14	6.5–7.5
6	8–20	4.5–5.5	5–14	5.5–6.5	8–30	6–7	12–24	5–6
7	8–30	5.5–6.5	8–30	6.5–7.5	8–20	5.5–6.5	12–24	4.5–5.5
8	8–20	6.5–7.5	8–30	6.5–7.5	8–20	5–6	5–14	6–7
9	8–30	5–6	12–24	5.5–6.5	8–20	5–6	5–14	5.5–6.5

In the user specific band, the feature extraction is performed by all three discussed feature extraction techniques, which are TDP, FFT, and PCA in the BCI system. As every individual subject has different frequency and timing bands, the dominant bands must also be identified for the individual subject. Accordingly, we select the frequency band and the timing using the Fisher ratio of the time domain parameter. Hence, a band pass filter is used to eliminate the insignificant bands and only obtain the significant sections during the Event Related Synchronisation (ERS) and Event Related De-synchronisation (ERD). Hence, the frequency range is selected between 5 and 30 Hz and the selected points are 5, 8, 12, 14, 20, 24, and 30.

For the subject specific band with FDA type, the F-score is computed for all of the ranges of the frequency–time region for identifying the optimal parameters and the maximum F-score. As the subject specific results are obtained, the optimal time–frequency area is identified for each individual test subject along with the subject–specific TDP features for training the classifier. Multiple class LDA is used for the classification for the subject specific results. The frequency range is selected between 8 and 30 Hz.

After performing both of them individually, they are combined to implement the proposed approach. For every subject, various time intervals are subjected to the three-feature extraction technique, which are FFT and TDP, and then feature reduction technique, which is PCA. After this, the classification techniques, Fisher's LDA and LS-SVM are performed. The number of frequency bands is calculated to identify which classification is optimum and which has a high accuracy. From the results, it is seen that the proposed combination of the feature extraction techniques works better when compared to its individual results. The frequency range for both the user-specific band and subject-specific band are identified and separated by applying the Discrete Fourier Transform (DFT) and Band Pass filter between the time interval. The analysis is statistically performed for both bands and the following points are selected as optimum, as shown in Table 2.

Table 2. Identified Optimum Intervals.

Subject	Time Interval
1	6.5
2	7.17
3	5.5
4	5.1
5	6.6
6	5.71
7	6.12
8	6.75
9	5.95

In order to identify the best time interval, it has to be individually identified for all the subjects. This is compared with the different frequency intervals for both the bands. Additionally, LDA is also performed for the same frequency intervals in order to identify the optimum number of layers and epochs. The different traditional feature extraction methods, like TDP, PCA, R^2 , FCBF, EMD, ITD, CV, and FFT, are compared with different combinations and tabulated in Table 3. The TDP and PCA are maintained as a constant technique, while modifying the other techniques— R^2 , FCBF, EMD, ITD, CV, and FFT along with Fisher's LDA as a classification. We are using the Fisher ratio of TDP. Therefore, TDP is constant. The highest average kappa coefficient accuracy obtained is 0.56 for the combination of TDP, FFT, and PCA. PCA is constant, because it is one of the most applied methods for feature reduction. From Table 3, it is observed that these combinations do not deliver high accuracy.

Table 3. Comparison of standard feature extraction methods.

Methods	Classification	1	2	3	4	5	6	7	8	9	Avg.
TDP + R^2 + PCA	Fisher's LDA	0.69	0.19	0.60	0.39	0.19	0.36	0.76	0.69	0.67	0.50
TDP + FCBF + PCA	Fisher's LDA	0.75	0.25	0.56	0.48	0.21	0.47	0.75	0.70	0.72	0.54
TDP + EMD + PCA	Fisher's LDA	0.65	0.18	0.65	0.38	0.16	0.30	0.69	0.66	0.65	0.48
TDP + ITD + PCA	Fisher's LDA	0.62	0.19	0.58	0.38	0.16	0.26	0.66	0.64	0.69	0.46
TDP + CV + PCA	Fisher's LDA	0.72	0.27	0.55	0.49	0.25	0.48	0.71	0.73	0.76	0.55
TDP + FFT + PCA	Fisher's LDA	0.74	0.26	0.57	0.48	0.27	0.49	0.70	0.75	0.78	0.56

The user specific band is performed with the combinations of TDP, FFT, and PCA with both LS-SVM and Fisher's LDA individually. For the LS-SVM classifier, the obtained average kappa coefficient accuracy of all nine subjects is seen to be 0.57, whereas, for the Fisher's LDA, the kappa coefficient accuracy is seen to be at 0.60. Currently, the FDA type F-score is performed with the combinations of TDP, FFT, and PCA with Fisher's LDA and the average kappa coefficient accuracy of 0.58 is obtained. The proposed technique is performed with the combination of TDP, FFT, and PCA, along with the user specific and FDA type F-score by using the Fisher's LDA classifier. With the proposed technique, it can be seen that the kappa coefficient accuracy is 0.64, which is

significantly higher than the traditional approaches. This comparison is tabulated in Table 4 and represented in Figure 8.

Table 4. Comparison of the proposed method with other combinations.

Methods	Classification	1	2	3	4	5	6	7	8	9	Avg
TDP + FFT + PCA (user specific band Fisher ratio)	LS-SVM	0.81	0.26	0.58	0.47	0.25	0.48	0.79	0.76	0.81	0.57
TDP + FFT + PCA (user specific band Fisher ratio)	Fisher's LDA	0.82	0.29	0.63	0.49	0.28	0.49	0.78	0.79	0.84	0.60
TDP + FFT + PCA (FDA type F-score)	Fisher's LDA	0.80	0.28	0.67	0.44	0.26	0.48	0.79	0.73	0.78	0.58
TDP + FFT + PCA (Combined both Fisher ratio and FDA type F-score)	Fisher's LDA	0.84	0.36	0.66	0.52	0.30	0.55	0.79	0.84	0.89	0.64

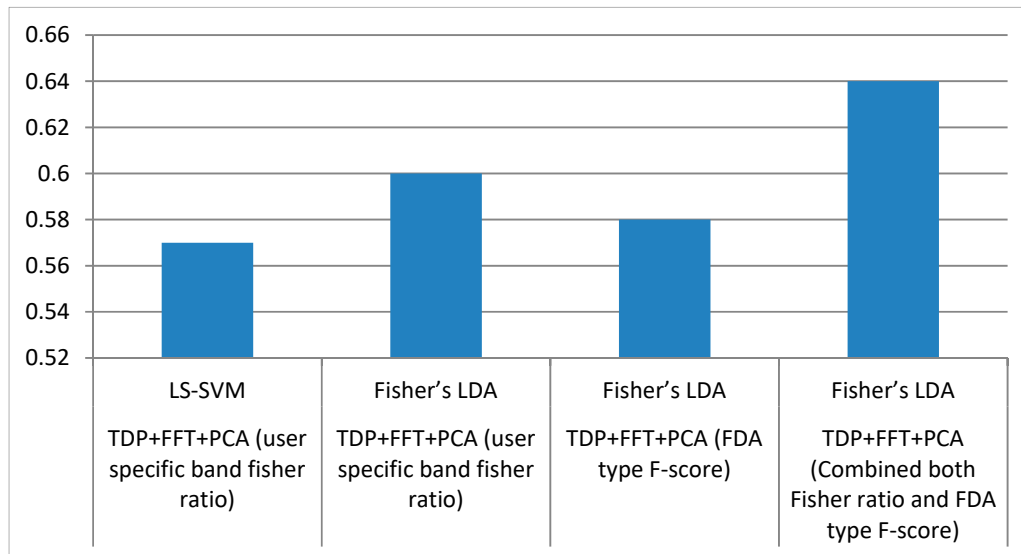


Figure 8. Comparison of the proposed method with other combinations.

Table 5 shows a comparison of the results of the proposed method and other competitive methods. The proposed technique of the user specific band is performed with combinations of TDP, FFT, and PCA (combined both Fisher ratio and FDA type F-score) with Fisher's LDA, and the average kappa coefficient accuracy is seen to be at 0.64. The combination of methods, such as FFT, CV, and PCA, are performed with the LS-SVM classifier, and the kappa coefficient accuracy is seen to be at 0.56 [9]. The FBCSP algorithm were used to optimize subject-specific frequency bands for the CSP algorithm to extract features and the average kappa coefficient accuracy is seen to be at 0.63 [10]. The combination of CNN and VAE methods achieved the kappa coefficient accuracy of 0.56 [18]. The FBCSP algorithm that employed the MIRS feature selection algorithm and yielded a kappa coefficient accuracy of 0.57 [31]. The Multiple discriminate analysis feature is performed with the SVM classifier and the kappa coefficient accuracy is seen to be at 0.55 [32]. The Autoregressive feature is performed with the LDA classifier, it can be seen that the kappa coefficient accuracy is 0.52 [33]. As a result, the proposed method offers a very satisfactory classification performance in comparison to the state-of-art methods. The obtained average result by the proposed method is 0.64,

which is higher than 0.63 [10], and it is much higher than the other results that were obtained on BCI competition data [Table 5]. Our proposed approach outperforms the CNN method [17, 18]. Thus, the combination of methods such as TDP, FFT, and PCA to the frequency bands selected by Fisher ratio and FDA type F-score along with Fisher's LDA yielded the most efficient results.

Table 5. Comparison of the results of the proposed method with the results of the state-of-art methods.

Subject	The Proposed Method	FFT + CV + PCA [9]	FBCSP [10]	CNN [17]	CNN + VAE [18]	FBCSP, MIRSR [31]	MDA [32]	AR [33]
1	0.84	0.80	0.82	-	0.52	0.40	0.71	0.62
2	0.36	0.28	0.51	-	0.34	0.20	0.31	0.44
3	0.66	0.55	0.80	-	0.43	0.21	0.75	0.65
4	0.52	0.49	0.51	-	0.90	0.95	0.47	0.42
5	0.30	0.30	0.37	-	0.64	0.85	0.19	0.40
6	0.55	0.41	0.33	-	0.64	0.61	0.20	0.35
7	0.79	0.79	0.83	-	0.55	0.55	0.78	0.51
8	0.84	0.68	0.78	-	0.50	0.85	0.77	0.64
9	0.89	0.77	0.72	-	0.51	0.74	0.73	0.64
AVG.	0.64	0.56	0.63	0.57	0.56	0.59	0.55	0.52

MIRSR: mutual information-based rough set reduction; MDA: Multiple discriminant analysis; AR: Autoregressive; for [17], the result was converted to kappa value.

5. Conclusions

In this paper, we present a novel method that is based on user specific band and subject specific band to select the frequency band and time segment for the classification of four-class motor imagery tasks. The method of common spatial patterns was used for pre-processing of the signal. The method reduced the number of channels from 22 to eight for the data set 2a from the BCI Competition IV. The different time intervals were examined with CSP, TDP, FFT, and PCA feature extraction methods and Fisher's linear discriminant analysis (LDA) classifier to find out the best time intervals each subject who performed motor imagery. We counted the number of frequency bands in which Fisher's linear discriminant analysis (LDA) classifier had the best accuracy rates in order to show that Fisher ratio and FDA-type F-score bands are effective in performance improvement. It is seen that, most of the time, a combination of Fisher ratio and FDA-type F-score bands would have better performance than just using Fisher ratio and FDA-type F-score bands alone. Different feature extraction techniques, like TDP, R², PCA, FCBF, EMD, ITD, CV, TDP, and FFT, are compared in this paper. From the different combination of the feature extraction techniques, it is observed that the combination of both user specific band and subject specific band improve the accuracy for the combination of time domain parameters, fast Fourier transform, and principal component analysis. The combination of these algorithms has significantly increased the accuracy when it is compared to the individual approaches. Different other combinations of the feature extraction techniques were also executed and compared to compare with the proposed approach. While the other standard combinations achieved the kappa coefficient accuracies of between 0.50 and 0.60, the proposed algorithm achieved a kappa coefficient accuracy of 0.64, which is significantly higher than the other approaches. The novelty of the approach is based on the combination of user specific band and subject specific band and two feature extraction methods. In the future, other combinations of feature extraction and classification approaches will be performed to further improve the accuracy.

Author Contributions: V.J. designed the task, supervised research, analyzed the results, provided feedback and gave new ideas, revised the draft and approved the final version of the article. S.G.S. implemented test software, executed experimental work, analyzed the results and did revisions to the final article.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Wolpaw, J.; Wolpaw, E.W. *Brain-Computer Interfaces Principles and Practice*; Oxford University Press: Oxford, UK, 2012.
2. Bansal, D.; Mahajan, R. *EEG-Based Brain-Computer Interfaces*. Elsevier: Amsterdam, The Netherlands, 2019; volume 4, pp. 21–71.
3. Martínez-Montes, E.; Valdés-Sosa, P.A.; Miwakeichi, F.; Goldman, R.I.; Cohen, M.S. Concurrent EEG/fMRI analysis by multiway Partial Least Squares. *Neuroimage* **2004**, *22*, 1023–1034.
4. DellaBadia, J., Jr.; Bell, W.L.; Keyes, J.W., Jr.; Mathews, V.P.; Glazier, S.S. Assessment and cost comparison of sleep-deprived EEG, MRI and PET in the prediction of surgical treatment for epilepsy. *Seizures* **2002**, *11*, 303–309.
5. Baig, M.Z.; Aslam, N.; Shum, H.P.H.; Zhang, L. Differential evolution algorithm as a tool for optimal feature subset selection in motor imagery EEG. *Expert Syst. Appl.* **2017**, *90*, 184–195.
6. García-Laencina, P.J.; Rodríguez-Bermudez, G.; Roca-Dorda, J. Exploring dimensionality reduction of EEG features in motor imagery task classification. *Expert Syst. Appl.* **2014**, *41*, 5285–5295.
7. Gupta, A.; Agrawal, R.K.; Kaur, B. Performance enhancement of mental task classification using EEG signal: A study of multivariate feature selection methods. *Soft Comput.* **2015**, *19*, 2799–2812.
8. Han, J.; Zhao, Y.; Sun, H.; Chen, J.; Ke, A.; Xu, G.; Zhang, H.; Zhou, J.; Wang, C. A Fast, Open EEG Classification Framework Based on Feature Compression and Channel Ranking. *Front. J. Neurosci.* **2018**, *12*, 217.
9. Jusas, V.; Samuvel, S.G. Classification of Motor Imagery Using Combination of Feature Extraction and Reduction Methods for Brain-Computer Interface. *Inf. Technol. Control* **2019**, *48*, 225–234.
10. Mahmoudi, M.; Shamsi, M. Multi-class EEG classification of motor imagery signal by finding optimal time segments and features using SNR-based mutual information. *Australas. Phys. Eng. Sci. Med.* **2018**, *41*, 957–972.
11. Ren, W.; Han, M.; Wang, J.; Wang, D.; Li, T. Efficient feature extraction framework for EEG signals classification. In *Processing of the Seventh International Conference on Intelligent Control and Information (ICICIP)*, Fujian, China, 1 December 2016; pp. 167–172.
12. Rodríguez-Bermúdez, G.; García-Laencina, P.; Roca-Dorda, J. Efficient Automatic Selection and Combination of EEG Features in Least Squares Classifiers for Motor Imagery Brain-Computer Interfaces. *Int. J. Neural Syst.* **2013**, *23*, 1350015.
13. Wang, J.; Feng, Z.; Lua, N.; Luo, J. Toward optimal feature and time segment selection by divergence method for EEG signals classification. *Comput. Biol. Med.* **2018**, *97*, 161–170.
14. Yuan, M.-D.; Feng, D.-Z.; Shi, Y.; Liu, W.-J. Dimensionality reduction by collaborative preserving Fisher discriminant analysis. *Neurocomputing* **2019**, *356*, 228–243.
15. Yu, X.; Chum, P.; Sim, K.-B. Analysis the effect of PCA for feature reduction in non-stationary EEG based motor imagery of BCI system. *Optik (Stuttg)* **2014**, *125*, 1498–1502.
16. Zhang, Y.; Ji, X.; Liu, B.; Huang, D.; Xie, F.; Zhang, Y. Combined feature extraction method for classification of EEG signals. *Neural Comput. Appl.* **2017**, *28*, 3153–3161.
17. Uktveris, T.; Jusas, V. Application of Convolutional Neural Networks to Four-Class Motor Imagery Classification Problem. *J. Inf. Technol. Control* **2017**, *46*, 260–273.
18. Dai, M.; Zheng, D.; Na, R.; Wang, S.; Zhang, S. EEG Classification of Motor Imagery Using a Novel Deep Learning Framework. *Sensors* **2019**, *19*, 551.
19. Ramoser, H.; Muller-Gerking, J.; Pfurtscheller, G. Optimal spatial filtering of single trial EEG during imagined hand movement. *IEEE Trans. Rehabil. Eng.* **2000**, *8*, 441–446.
20. Wang, Y.; Gao, S.; Gao, X. Common spatial pattern method for channel selection in motor imagery based BCI. *Eng. Med. Biol.* **2005**, *5*, 5392–5395.
21. Oh, S.-H.; Lee, Y.-R.; Kim, H.-N. A novel EEG feature extraction method using hjorth parameter. *Int. J. Electron. Electr. Eng.* **2014**, *2*, 217.
22. Yang, Y.; Chevallier, S.; Wiart, J.; Bloch, I. Subject-specific time-frequency selection for multi-class motorimagery-based BCIs using few Laplacian EEG channels. *Biomed. Signal Proc. Control* **2017**, *38*, 302–311.
23. Vidaurre, C.; Krämer, N.; Blankertz, B.; Schlögl, A. Time Domain Parameters as a feature for EEG-based Brain-Computer Interfaces. *Neural Netw.* **2009**, *22*, 1313–1319.

24. Karatoprak, E.; Seker, S. An Improved Empirical Mode Decomposition Method Using Variable Window Median Filter for Early Fault Detection in Electric Motors. *Math. Problem Eng.* **2019**, *2019*, 1–9.
25. Senliol, B.; Gulgezen, G.; Yu, L.; Cataltepe, Z. Fast Correlation Based Filter (FCBF) with a different search strategy. In Proceedings of the 23rd International Symposium on Computer and Information Sciences, Binghamton, NY, USA, 29 November 2008, 1–4.
26. Frei, M.; Osorio, I. Intrinsic time-scale decomposition: Time–frequency–energy analysis and real-time filtering of non-stationary signals. *Proc. R. Soc. A Math. Phys. Eng. Sci.* **2007**, *463*, 321–342.
27. Vega, R.; Sajed, T.; Mathewson, K.W.; Khare, K.; Pilarski, P.M.; Greiner, R.; Sanchez-Ante, G.; Antelis, J.M. Assessment of feature selection and classification methods for recognizing motor imagery tasks from electroencephalographic signals. *Artif. Intell. Res.* **2017**, *6*, 37–51.
28. Jafarzadegan, M.; Safi-Esfahani, F.; Beheshti, Z. Combining hierarchical clustering approaches using the PCA method. *Expert Syst. Appl.* **2019**, *137*, 1–10.
29. Wen, X.; Tu, C.; Wu, M.; Jiang, X. Fast ranking nodes importance in complex networks based on LS-SVM method. *Phys. A Stat. Mech. Appl.* **2018**, *506*, 11–23.
30. Han, X.; Wang, J.; Wu, Z.; Li, G.; Wu, Y.; Li, J. Learning solutions to two dimensional electromagnetic equations using LS-SVM. *Neurocomputing* **2018**, *317*, 15–27.
31. Ang, K.K.; Chin, Z.Y.; Wang, C.; Guan, C.; Zhang, H. Filter bank common spatial pattern algorithm on BCI competition IV datasets 2a and 2b. *Front. Neuroscience* **2012**, *6*, 39.
32. Suk, H.-I.; Lee, S.-W. Subject and class specific frequency bands selection for multiclass motor imagery classification. *Int. J. Imaging Syst. Technol.* **2011**, *21*, 123–130.
33. Brunner, C.; Billinger, M.; Vidaurre, C.; Neuper, C. A comparison of univariate, vector, bilinear autoregressive, and band power features for brain-computer interfaces. *Med. Biol. Eng. Comput.* **2011**, *49*, 1337–1346.



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).