Research Paper

Classification of Right/Left Hand Motor Imagery by Effective Connectivity Based on Transfer Entropy in Electroencephalogram Signal

Erfan Rezaei[1], **Ahmad Shalbaf**[1*]

1. Department of Medical Physics and Biomedical Engineering, School of Medicine, Shahid Beheshti University of Medical Sciences, Tehran, Iran.

Use your device to scan and read the article online

Citation Rezaei, E., & Shalbaf, A. (2023). Classification of Right/Left Hand Motor Imagery by Effective Connectivity Based on Transfer Entropy in Electroencephalogram Signal. *Basic and Clinical Neuroscience, 14*(2), 213-224. http://dx.doi.org/10.32598/bcn.2021.2034.3

doi: **http://dx.doi.org/10.32598/bcn.2021.2034.3**

Article info:
Received: 08 Apr 2021
First Revision: 18 Jul 2021
Accepted: 18 Sep 2021
Available Online: 01 Mar 2023

Keywords:
Electroencephalogram (EEG), Motor imagery, Effective connectivity, Transfer entropy

ABSTRACT

Introduction: The right and left-hand motor imagery (MI) analysis based on the electroencephalogram (EEG) signal can directly link the central nervous system to a computer or a device. This study aims to identify a set of robust and nonlinear effective brain connectivity features quantified by transfer entropy (TE) to characterize the relationship between brain regions from EEG signals and create a hierarchical feature selection and classification for discrimination of right and lefthand MI tasks.

Methods: TE is calculated among EEG channels as the distinctive, effective connectivity features. TE is a model-free method that can measure nonlinear effective connectivity and analyze multivariate dependent directed information flow among neural EEG channels. Then four feature subset selection methods namely relief-F, Fisher, Laplacian, and local learningbased clustering (LLCFS) algorithms are used to choose the most significant effective connectivity features and reduce redundant information. Finally, support vector machine (SVM) and linear discriminant analysis (LDA) methods are used for classification.

Results: Results show that the best performance in 29 healthy subjects and 60 trials is achieved using the TE method via the Relief-F algorithm as feature selection and support vector machine (SVM) classification with 91.02% accuracy.

Conclusion: The TE index and a hierarchical feature selection and classification can be useful for the discrimination of right- and left-hand MI tasks from multichannel EEG signals.

*** Corresponding Author:**
Ahmad Shalbaf
Address: *Department of Medical Physics and Biomedical Engineering, School of Medicine, Shahid Beheshti University of Medical Sciences, Tehran, Iran.*
Tel: *+98 (912) 2163201*
E-mail: *shalbaf@sbmu.ac.ir*

Highlights

• Effective connectivity features were extracted from electroencephalogram (EEG) to analyze relationships between regions.

• Four feature selection methods used to select most significant effective features.

• Support vector machine (SVM) used for discrimination of right and left hand motor imagery (MI) task.

Plain Language Summary

In this study, we investigated brain activity using effective connectivity during MI task based on EEG signals. The motor imagery task can accomplish the same goal as motor execution, since they are both activated by the same brain area. Transfer entropy, coherence, and Granger casualty were employed to extract the features. Differential patterns of activity between the left vs. right MI task showed activity around the motor area rather than other areas. In order to reduce redundant information and select the most significant effective connectivity features, four feature subset selection algorithms are used: Relief-F, Fisher, Laplacian, and learning-based clustering feature selection (LLCFS). Then, support vector machine (SVM) and linear discriminant analysis (LDA) are used to classify left and right hand MI task. Comparison of three different connectivity methods showed that TE index had the highest classification accuracy, and could be useful for the discrimination of right and left hand MI tasks from multichannel EEG signals.

1. Introduction

Brain-computer interface (BCI) is a method (Matthews et al., 2007; Yin et al., 2015) that helps paralyzed people to communicate with the environment without using the peripheral nervous system and muscles (Wolpaw et al., 2002). Electroencephalography (EEG) (Wolpaw et al., 2000), magnetoencephalography (MEG) (Mellinger et al., 2007), fMRI (Sitaram et al., 2007), and near-infrared spectroscopy (NIRS) (Naseer & Hong, 2015) are non-invasive methods for BCI system. Among them, for practical applications, EEG, which measures brain activity via metal electrodes positioned on the scalp, is the most used due to its noninvasive, low cost, and high temporal resolution and is widely used in neuroscience applications (Cincotti et al., 2008; Rehan & Hong, 2012). A type of BCI known as motor imagery (MI) refers to the imagination of particular action without actual execution. Instead of doing another mental task or multiple control command, MI offers a more efficient approach for healthy people to learn new skills and paralyze people to rehabilitate (Sharma et al., 2006). MI studies usually use different limbs' imagination, such as the left and right hand, feet, and tongue.

In the last decades, several studies from one-channel EEG have been presented to classify MI. Power spectral density (Kim et al., 2018), discrete wavelet transform (Athif & Ren, 2019), autoregressive model coefficients (Jansen et al., 1981), common spatial pattern (CSP) (Park et al., 2017; Shin et al., 2016), sparse representation (Shin et al., 2012), and Hilbert transform (Sun et al., 2015) were employed for feature extraction from EEG signals. Despite significant results, none of these methods have been proven to be adequately reliable in practical settings because features from single-channel EEG during MI tasks cannot attain reliable information and multi-channel EEG-based features must be measured.

The functional and effective brain connectivity analyses are powerful tools to investigate the relationship among different brain regions for the EEG analysis to identify the complicated neurophysiological changes during task performance (19-20). Functional connectivity is generally inferred by the correlation, coherence, phase lag index (PLI), and phase lock value (PLV) in EEG signals (Qin et al., 2010; Rubinov & Sporns, 2010; Vinck et al., 2011). PLV was calculated for event-related desynchronization/synchronization (ERD/S) between two types of MI tasks (Gu et al., 2020) and for the phase coupling of sensorimotor during tounge-MI task (Spiegler et al., 2004). Santamaria and James employed PLV and wavelet coherence to classify two different MI tasks with six different classification algorithms (Santamaria & James, 2016). PLV was also used directly to categorize left and right hand MI (Gonuguntla et al., 2016) and during left and right hand, foot, and tongue imagery (Brunner et al., 2006). Finally, Hamedi et al. used coherence to classify four distinct MI tasks (Hamedi et al., 2015).

Another popular brain connectivity widely used in neuroscience is effective connectivity which refers to the influence that one neural system exerts over another (Friston, 2011). These methods are used to define directional effects between any pair of EEG signals. Effective connectivity is usually computed using several methods, including structural equation modeling (SEM) (Anderson & Gerbing, 1988), dynamic causal modeling (DCM) (Harrison & Penny, 2003), and Granger causality (GC) (Granger, 1988). However, SEM is not appropriate for time series, in which the time constants of neuronal hemodynamics are generally much larger than the fluctuating inputs driving them (Friston, 2011). Moreover, DCM, which partially covers nonlinear interactions, always requires selecting a prior model in advance. In other words, a priori information about system input and connections between the system's parts is required (Friston et al., 2003). However, this information always is not available. Finally, GC, partial directed coherence (PDC), and directed transfer function (DTF) use a linear stochastic model for the signal's intrinsic dynamic and limit the effective connectivity pattern to specific templates (Chen et al., 2019; Granger, 1969; Liang et al., 2016; Rathee et al., 2016). PDC and DTF can determine the direction and spectral characteristics of the EEG signals simultaneously (Kaminski & Blinowska, 1991; Sameshima & Baccalá, 1999). Liang et al. employed PDC combined with multivariate empirical mode decomposition during the left/right hand MI tasks to improve classification performance (Liang et al., 2016). DTF has been used to investigate brain activity dynamics (Kamiński et al., 2001) and evaluate motor task experiments (Ginter et al., 2001). The pattern of EEG in beta and gamma bands during left and right-hand movement imagination was also investigated by DTF (Ginter et al., 2005; Kus et al., 2004). Generalized partial directed coherence (GPDC), partially directed coherence factor (PDCF), and full-frequency DTF (ffDTF) are other extensions of the effective connectivity methods that have been used in the BCI (Billinger et al., 2013).

In summary, these effective connectivity methods have these problems, the need for a priori information or model, the inability to detect nonlinear connections, the inability to detect all connectivity in the complex network, and the lack of robustness against linear cross-talk between electrophysiological signals (Nalatore et al., 2007; Nolte et al., 2008). Therefore, characterizing and understanding brain dynamics during MI tasks should be done in a way that does not have the mentioned problems. Hence, an essential nonlinear criterion to estimate effective connectivity with the name of transfer entropy (TE) is presented (Schreiber, 2000). In this method, the Wiener causality concept and the conditional mutual information in the context of information theory are combined. This method is a model-free method and does not need a priori assumptions on connectivity patterns due to its exploratory nature, robustness against linear crosstalk, and can measure all linear and nonlinear effective connectivity between brain regions. This method recently has become popular and widely applied to analyze multi-channel EEG signals (Weinrich & Wise, 1982).

After the EEG feature extraction, these features must be optimized by a wide variety of feature selections or dimensional reduction, such as principal component analysis (Vickers, 2017), independent component analysis (Ruan et al., 2018), and sequential floating forward search (Asensio-Cubero et al., 2016). Finally, these optimized features must be classified to develop MI tasks, such as Bayesian classifier (He et al., 2015), extreme learning machine (Hsu, 2015), and deep learning approaches (Zhang et al., 2019). Despite the different machine learning algorithms, no universally superior algorithm exists for this application.

This study aims to provide a nonlinear effective connectivity method named TE index as feature extraction to characterize the nonlinear directed interaction among neural EEG channels. Four different feature selection methods and advanced classification methods are used to improve the accuracy of discrimination of left vs. right hand MI tasks in 29 participants.

2. Materials and Methods

Participants and experimental design

Twenty-nine healthy subjects (fourteen males and fifteen females with a mean of 28.5±3.7 years) with no reported brain-related diseases participated in this study (Shin et al., 2016). The subjects sat on a comfortable chair at a 1.6-meter distance from a 50-inch white screen and were not allowed to move their body during the task. The experiment consisted of three sessions of right and left-hand MI. Each session consisted of 60 s of rest before the experiment, 20 repetitions of the task (10 trials for each left and right hand MI), and 60 s post-experiment resting period. After the task started with a visual introduction, a task period of 10 s, and a resting time between 15 to 17 s, which is randomly done. In the task period, subjects imagine opening and closing their hands at a speed of 1 Hz. Figure 1 shows the schematic diagram of the experiment. Therefore, for each subject in the whole three sessions, 30 trials for left and 30 trials for right hand MI were performed. EEG raw data was recorded at a 1000 Hz sampling rate and then downsampled to 200 Hz. EEG electrodes were placed at the same cap according to the international 10-5 system with thirty active electrodes (AFp1, AFp2, AFF1h, AFF2h, AFF5h, AFF6h, F3, F4, F7, F8, FCC3h, FCC4h, FCC5h, FCC6h, T7, T8, Cz, CCP3h, CCP4h, CCP5h, CCP6h, Pz, P3, P4, P7, P8, PPO1h, PPO2h, POO1, POO2, and Fz as ground and reference electrode).

Preprocessing

All data pre-processing was done using EEGLab in MATLAB sotware, version R2018b. The measured EEG data were filtered by a 1 Hz FIR high pass filter and re-referenced using a common average reference. Independent component analysis (ICA)-based electro-oculogram (EOG) rejection was performed by toolbox in EEGLAB (Delorme & Makeig, 2004).

Effective connectivity

Effective connectivity as a feature extraction method is used to study brain communication mechanisms of different areas (direction and strength of the information). Effective connectivity is used to analyze more than one signal simultaneously and refers to a casual activity that one neural network has on the activity of another neural network (Friston, 2011). In this study, TE was used to estimate nonlinear effective connectivity. All effective connectivity calculation was done in MATLAB.

Transfer entropy (TE)

TE is a nonparametric method for estimating effective connectivity, that can measure all linear and nonlinear causal relationships (Vicente et al., 2011). In this method, the Wiener causality concept and conditional mutual information are combined. MI(x, y) shows the mutual information and is defined by Equation 1.

$$1.\ MI(x,y)=\sum_{(x,y)} p_{(x,y)} \log\left(\frac{p(x.y)}{p(x)p(y)}\right)$$

In Equation 1, p(x) and p(y) are the probability density functions, and p(x,y) is the joint probability density function. MI(x, y) can be rewritten using Shannon entropy according to Equation 2 (Vicente et al., 2011).

$$2.\ MI(x.y)=H(x)+H(y)-H(x.y)=H(x)-H(x|y)=H(y)-H(y|x)$$

In Equation 2, H(x) is Shannon entropy, H(x, y) is the joint entropy of random variables. Also, H(x|y) and H(y|x) are the conditional entropies (Vicente et al., 2011):

Conditional mutual information MI (x.y|z) is dependent on observing the random variable z and is calculated by Equations 3 and 4 (Vicente et al., 2011).

$$3.\ MI(x.y|z)=\sum_{x.y.z} p(x.y.z)\log\left(\frac{p(x.y|z)}{p(x|z)p(y|z)}\right)=\sum_{x.y.z} p(x.y.z)\log\left(\left(p\left(\frac{p(x.y.z)p(z)}{p(x.z)p(y.z)}\right)\right.\right.$$

$$4.\ MI(x.y|z)=H(x.z)+H(y.z)-H(z)-H(x.y.z)$$

By combining the Wiener causality and the MI (x.y|z), TE is obtained (Vicente et al., 2011). or TE(x→y) or TE_{xy} expressions that by assumption of knowing the past statement of the random variable x, how much it adds to the available information about the random variable y (Vicente et al., 2011).

$$5.\ TE(x\rightarrow y)=TE_{xy}=MI(y(t+\tau), \underline{x}_t^{dx,\tau_x} \mid \underline{y}_t^{dy,\tau_y})$$

$$\underline{x}_t^{dx,\tau_x}=(x(t).x(t-\tau\_x). \dots .x(t-(d\_X-1)\ \tau\_X))$$

$$\underline{y}_t^{dy,\tau_y}=(y(t).y(t-\tau\_y). \dots .y(t-(d\_y-1)\ \tau\_y))$$

(x_t^{dx,τ_x}) and ($\underline{y}_t^{dy,\tau_y}$) are the past status vectors.τ_X and τ_Y are embedding delay, and x and y and d_X and d_Y are embedding dimensions of x and y, respectively. When TE=0, no causality is observed between x and y and (>0): x is causing y. Embedding dimension (d) is the memory of the Markov process in each signal. Also, embedding delay (τ) is the autocorrelation time of the signal, i.e. when the envelope of the autocorrelation function decreases to 1/e (0.32).

Feature selection

The insignificant extracted features obtained from effective connectivity methods must be deleted. This reduction of features and selection of the best features can influence the improvement of classification performance. Feature selection methods are divided into three models, wrapper, embedded, and filter methods. The wrapper method employs classifiers to score a given subset of features, and the embedded method utilizes a selection process to learn classifiers. In contrast, the filter selection methods are based on the general characteristics of data, and any predictor and classifiers are ignored (Roffo, 2016). In the filter selection method, which is based on the intrinsic properties of data, features are considered individually, ranked, and then a subset is extracted. In this study, four widely filter selection algorithms named, relief-F, Fisher, Laplacian, and local learning-based clustering feature selection (LLCFS) algorithms are employed to choose the best features.

Relief-F is an iterative, randomized, and supervised method (Liu & Motoda, 2008). Fisher, which is a supervised method, calculates a feature score as the ratio of interclass separation and intra-class variance, where the feature is evaluated independently (Gu et al., 2012). In the Laplacian score, an unsupervised method, its power of locality preservation evaluates the importance of a feature based on the nearest neighbor graph (He et al., 2005). Finally, the LLCFS algorithm tries to ensure that the cluster tag of each data point is near the tag predicted by the local regression model, with its adjacent points and their cluster tags (60).

Classification

Support vector machine (SVM) (Meyer et al., 2003) and linear discriminant analysis (LDA) (Tharwat et al., 2017) as supervised learning algorithms are used for classification in this study. The LDA method is used to find a linear combination of features. SVM is used as a binary classifier to categorize the data to maximize the margin between the hyperplane and the nearest data. In this method, when overlapped features exist, support vector classification maps the feature into a higher dimension space by nonlinear function and creates an excellent discriminatory hyperplane in that space. All analyses were computed in MATLAB software, version 2018 (The Mathworks, Inc., Natick, MA, USA).

Statistical analysis

Ten-fold cross-validation was used in this study due to the limited dataset. In this method, data are divided into 10 parts of equal sizes, and in each run, the classification parameters are constructed with 90% of data (80% train and 10% validation [for selecting the optimized number of features]) and tested with 10%of data. Therefore, in the first step, we used only 90% of the data, and 10% of the test data is set aside in each run. When the optimal number of features is selected based on the validation dataset for 10 folds, the final classification results are reported based on the testing data of each 10 fold. Evaluation performance is reported by averaging the ten results.

3. Results

We calculated the effective and functional connectivity between all EEG signals using TE, coherence, and Granger causality index in 10-s windows in each trial run for the whole experimental period and all subjects. For 29 subjects in the entire three sessions with performing 30 trial runs for left and 30 trials for right hand MI, we have 30×29=870 trial runs for each class in the classification procedure. We used the Hermes toolbox to extract features. Having a 30-channel EEG, 900(30×30) connections between channels as functional and effective connectivity features are extracted, making further computations complex. As a result, we performed four feature selection methods (Fisher, LLCFS, Laplacian, and relief-F) to choose the best features for the discrimination of left and right MI tasks. Finally, the best-selected features are fed to SVM and LDA classifiers to classify EEG data into left vs. right MI tasks in 29 participants. We evaluated different kernels and different parameters in validation data through trial and error and finally used an SVM with RBF kernel and sigma of 0.9. A-10 fold cross-validation was performed. In our case, we used 10% of the data for tests and 90% for training and validation (80% train and 10% validation). Figure 2 shows the diagram of the proposed method.

Table presents the testing classification accuracies obtained by the TE, coherence and GC measure and four feature selection methods and classification methods over all participants. As can be seen, the proposed method by TE revealed better results rather than other connectivity methods (coherence and GC). Also, as can be seen, the SVM classifier revealed better results rather than LDA in all feature selection methods. Finally, TE with SVM and feature selection via the relief-F method yield the best results with high testing classification accuracy (91.02%). It is noteworthy that the best testing classification accuracy was obtained in the TE method with a smaller number of features rather than in GC and coherence methods. Afterward, TE with SVM and feature selection via Fisher's method yields a testing classification accuracy of 86.93. Figure 3 shows raw 900(30×30) connectivity features for the TE method overall left vs. right MI tasks. In this Figure, a higher absolute value of the connectivity feature is shown in warmer colors. We have also plotted P of all connectivity features by the TE method between the left and right hand to show the map of separability in Figure 4. P for 30×30=900 features (except 30 diagonal channels) between two classes are calculated and plotted. A lower value of P, shown in blue, has more separability. As can be seen, during the MI task, EEG signals have high separability around the motor area, parietal, and temporal. Figure 5 shows the results of validation classification accuracy by SVM vs. the number of TE features selected by relief-F and Fisher, which have the highest classification accuracy. Each box plot demonstrates the results of validation accuracy for each number of features for 10 folds. As can be seen, with an increasing number of selected features, the accuracy of the classification reaches its maximum value and then decreases. In these figures, the numbers 12 and

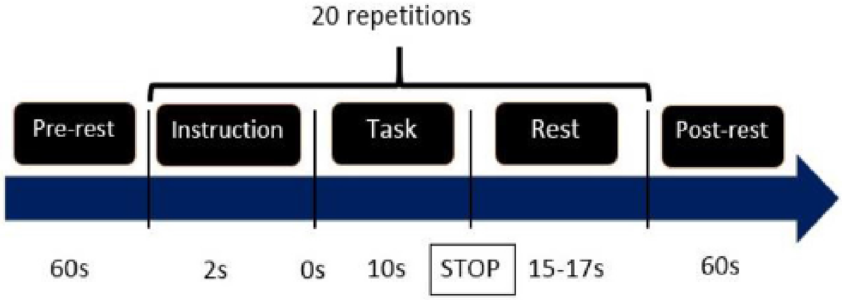


Figure 1. Schematic diagram of the experimental paradigm NEUROSCIENCE

13 of the features have high average validation classification accuracy for 10 folds. When the optimal number of features is selected based on the validation dataset, the final classification results are reported in Table 1 based on the testing data.

4. Discussion

In this article, a new automated method based on nonlinear effective connectivity named TE index among different brain regions as features and hierarchical machine learning algorithms is used to discriminate of left vs. right hand MI task from 30-channel of EEG signals in 29 participants with a satisfactory testing classification accuracy of 91.02%. TE index, which measures the transfer of information between collaborative processes based on information theory, is a suitable index during hand MI tasks. The novelties of our paper are the use of the TE method to quantify the connectivity of the EEG signals during the MI task and proposing a hierarchical machine learning structure based on different feature selection methods (relief-F, lapalacian, LLCFS, and Fisher) to filter the best discriminative features and then fed to SVM classifier.

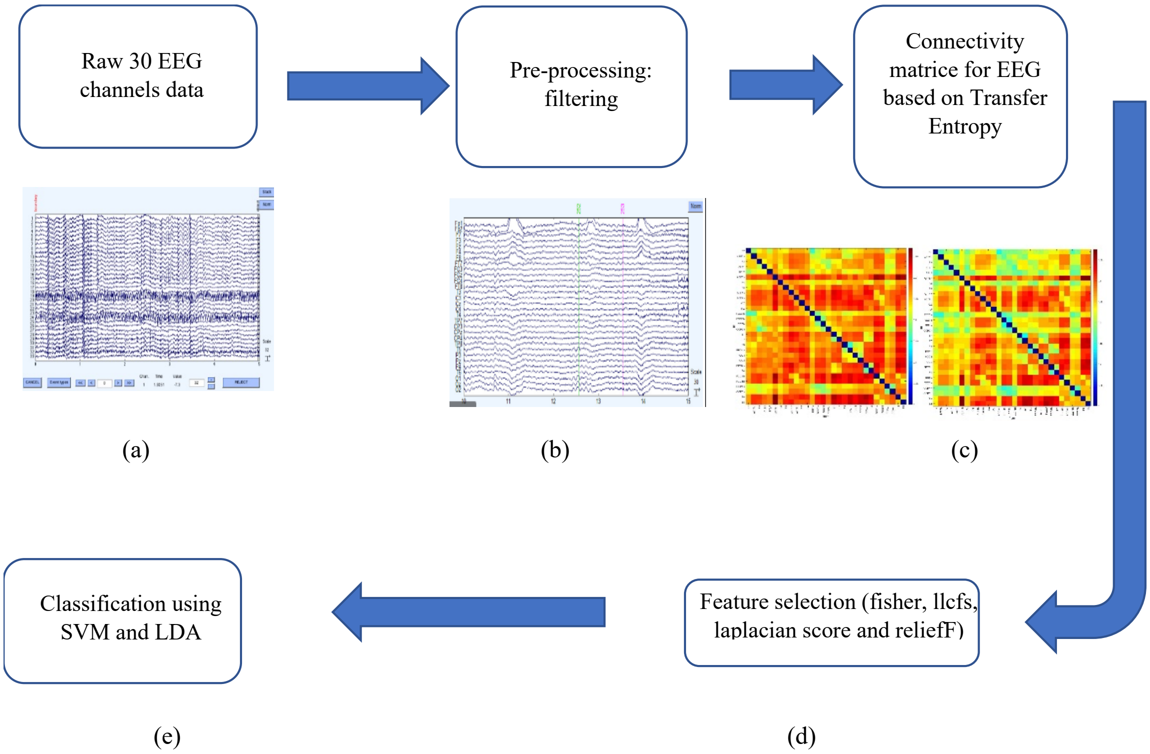
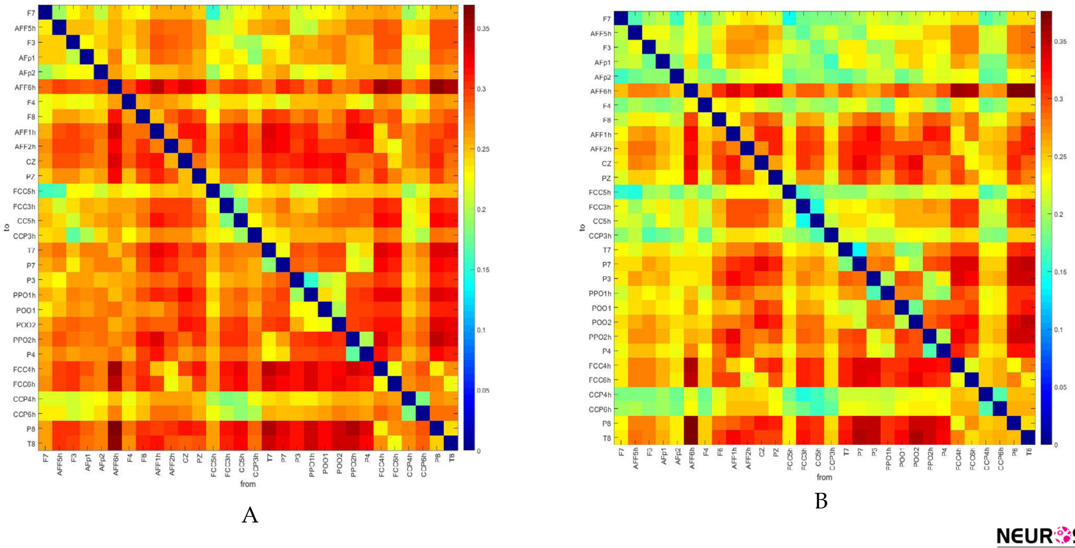


Figure 2. The process of the proposed system, raw EEG data NEUROSCIENCE

A) Preprocessing; B) Construction of effective connectivity features; C) Selection of significant extracted connectivity features using relief-F, Fisher, Laplacian, and LLCFS and ranking them; D) Classification using SVM and LDA.

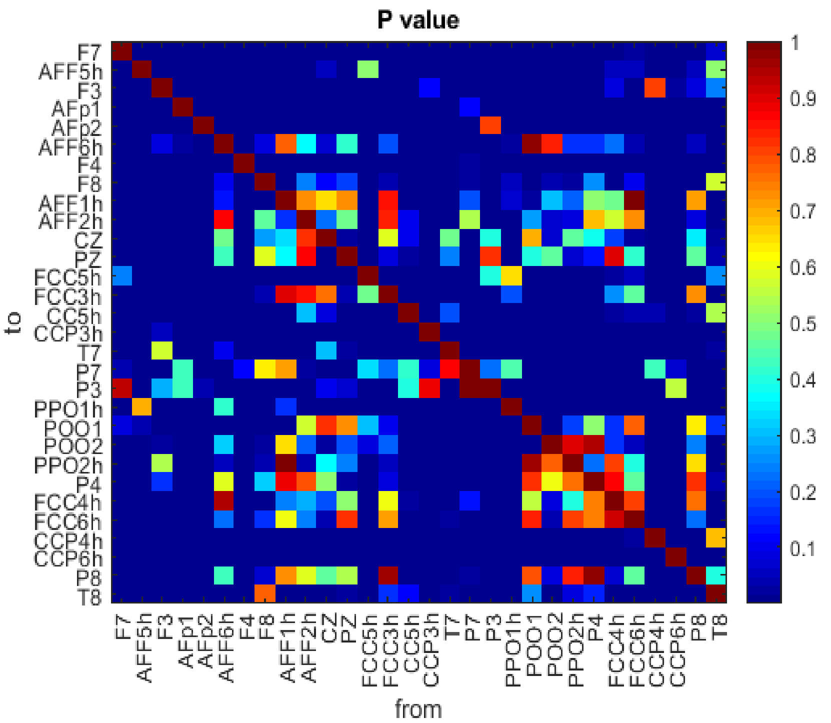


NEUROSCIENCE

Figure 3. Raw 900(30×30) TE connectivity features overall participants for MI task, a higher absolute value of connectivity feature showing with warm colors

A) Transfer entropy EEG features for right hand imagination; B) Transfer entropy EEG Features for left hand imagination.

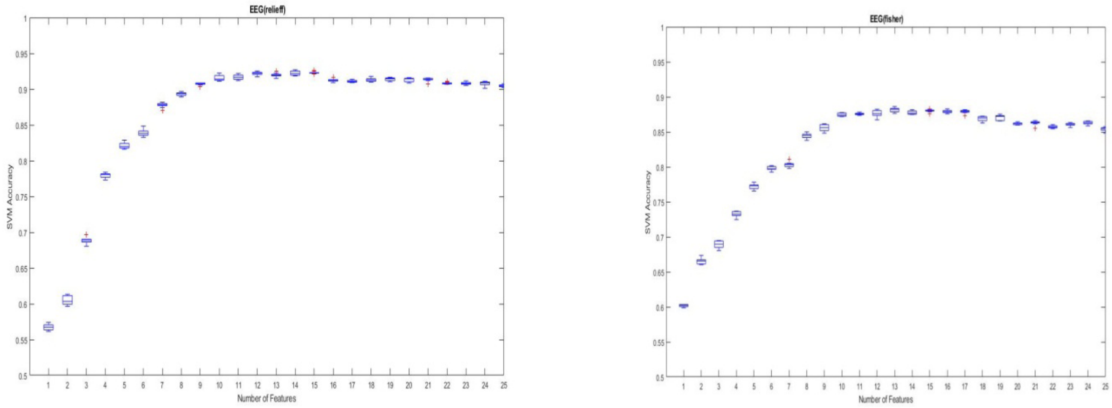
To estimate effective brain connectivity for the left vs. right hand MI task, we had several options, such as SEM, GC, and DCM. However, each of these methods has some problems. SEM is not appropriate for time series, in which the characteristic time constants of neuronal hemodynamics are much larger than the fluctuating or exogenous inputs that drive them. GC methods limit the effective pattern to specific templates based on the linear parameters of the MVAR model. In contrast, the natural dynamics of brain connections cannot be simply



NEUROSCIENCE

Figure 4. P of all connectivity features by TE method between left and right hand to show the map of separability

A lower value of P is shown in blue, has more separability.



NEUROSCIENCE

Figure 5. SVM validation classification accuracy vs. the number of TE features selected by (left) Fisher and (right) relief-F feature selection methods

isolated with a predetermined limiting model. Finally, DCM, which partially covers nonlinear interactions, always requires a priori information about the network connections, and this information is not always available. Consequently, if we want to use an appropriate effective connectivity method for a complex network of the brain, it should not require a priori information, be able to detect and measure nonlinear interaction across brain function, be robust against linear cross-talk between signals because EEG contains electrophysiological data, and finally detect effective connectivity with a wide distribution of interaction, because signaling between two areas of the brain may involve various pathway over various axons that connect two areas. Therefore, the use of the above-mentioned effective connectivity methods leads to incorrect brain connection estimation, and an important nonparametric and nonlinear criterion for estimating effective connectivity with the name of TE for studying multi-channel EEG signals is presented. TE index does not assume any particular model and can describe linear and nonlinear interactions existing in a system quantitatively and can properly detect directional connectivity.

Motor imagery and motor execution activate the same brain area (Beisteiner et al., 1995), and the same goal can be achieved by motor imagery task. During the interval between the motor imagery and the motor execution, several cells in the premotor cortex fired vigorously and then stopped firing after the execution (Weinrich & Wise, 1982). As shown in Figures 3, and 4, differential patterns of effective connectivity between the left vs. right MI task in the TE method are around the motor areas rather

Table 1. The classification accuracy obtained by TE, coherence, and GC measures

Connectivity Mthod	Classifier	Feature Selection Methods (No.)			
		Relief-F	Laplacian	LLCFS	Fisher
TE	LDA	85.74±0.002 (14)	82.31±0.0021 (18)	71.69±0.0011 (15)	84.9±0.002 (14)
	SVM	91.02±0.0015 (12)	86.87±0.0031 (17)	85.21±0.0030 (13)	86.93±0.003 (13)
Coherence	LDA	75.12±0.002 (20)	66.32±0.002 (25)	67.85±0.001 (19)	63.45±0.002 (22)
	SVM	80.67±0.0021 (17)	81.42±0.003 (23)	79.69±0.003 (17)	67.42±0.003 (23)
GC	LDA	80.13±0.0012 (18)	82.36±0.0036 (20)	74.63±0.0021 (18)	73.35±0.002 (17)
	SVM	82.48±0.0021 (16)	83.64±0.0026 (18)	84.12±0.0029 (17)	74.36±0.0034 (15)

NEUROSCIENCE

Abbreviations: TE: Transfer entropy; LLCFS: Local learning based clustering; SVM: Support vector machine; LDA: Linear discriminant analysis; GC: Granger causality; LLCFS: Local learning-based clustering feature selection.

Table 2. Comparison our work results with the prior studies employing the same data set inEEG Signal

Authors	Dataset	Year	Feature	Classifier	Accuracy
Shin (Shin et al., 2016)	29 subject, Shin dataset	2016	CSP	LDA	65.6
Yavuz (Yavuz & Aydemir, 2017)	29 subject, Shin dataset	2017	Hilbert transform	Knn	82.23
				LDA	78.13
Masghsoud I (Maghsoudi & Shalbaf, 2022)	29 subject, Shin dataset	2020	DTF	SVM	61.85
			dDTF		73.66
			GPDC		83.87
Proposed method	29 subject, Shin dataset	2021	TE	SVM	91.02
				LDA	85.74

NEUROSCIENCE

Abbreviations: TE: Transfer entropy; SVM: Support vector machine; LDA: Linear discriminant analysis; CSP: Common spatial pattern; DTF: Directed transfer function; GPDC: Generalized partial directed coherence; dDTF: Direct directed transfer function; Knn: K-nearest neighbors.

than other areas. These discriminative features are similar to previous studies (Hermes et al., 2011). Also, the parietal and temporal area has more activity than other areas, resulting from imagination, and the frontal area has some activity due to the decision thinking during the MI task (Xu et al., 2014).

In this study, four widely featured selection methods (relief-F, Fisher, Laplacian, and LLCFS) were applied. In these methods, features are individually ranked, and then a feature subset is extracted for classification in the next step. Relief-F and then Fisher feature selection methods which are supervised methods yielded better classification results in our study. The relief-F method, with the highest accuracy, considers the relevance of features with dependent variables using statistical measures and estimates the quality of the features. Also, K-nearest neighbors are searched and their contribution to each feature's weight is averaged to prevent redundant and noisy features that affect the nearest neighbors' selection.

Table 2 compares the results of our work with the prior studies that employed the same data set in EEG signal to classify left vs. right hand MI task (Shin et al., 2016; Yavuz & Aydemir, 2017). As it is observed, the testing accuracy achieved in this study by applying the effective connectivity method with TE as the features and hierarchical feature selection and SVM classifier (91.02%) is higher than those studies. It proves the preference for the proposed method. In the future, it is suggested to calculate the characteristics of the effective brain connectivity in the localization of brain resources in the EEG signal with more channels and then discuss the features and classification. Also, we believe that the performance of a multi-model system based on EEG and near-infrared spectroscopy using Hbo and HHB as the hemodynamic responses (Erdoğan et al., 2014; Gagnon et al., 2014) compared to a single modality may improve the accuracy of hand MI task discrimination.

5. Conclusion

Results indicated that nonlinear effective connectivity between brain regions using TE with handling other problems of previous effective connectivity characterizes brain dynamics effectively, and is an essential tool for understanding the neurophysiological mechanisms of left vs. right hand MI task. Consequently, by calculating the features of effective connectivity quantified with TE and a feature selection and SVM classifier to discriminate left vs. right MI task from EEG signals, a test accuracy of 91.02% on the 29 participants is achieved.

Ethical Considerations

Compliance with ethical guidelines

All procedures performed in studies involving human participants were in accordance with the declaration of Helsinki and was approved by the Ethics Committee of the Institute of Psychology and Ergonomics, Technical University of Berlin (Code: SH_01_20150330).

Funding

The present article is financially supported by the "Research Department of the School of Medicine Shahid Beheshti University of Medical Sciences" (Grant No.: 20267).

Authors' contributions

Conceptualization, methodology, investigation, writing- reviewing and editing, and validation: All authors; Supervision: Ahmad Shalbaf; Writing- original draft preparation: Erfan rezaei, Ahmad Shalbaf.

Conflict of interest

The authors declared no conflict of interest.

Acknowledgments

The authors would like to thank Research Department of the School of Medicine Shahid Beheshti University of Medical Sciences for the support.

References

Anderson, J.C. & Gerbing. D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin, 103*(3), 411-423. [DOI:10.1037/0033-2909.103.3.411]

Asensio-Cubero, J., Gan, J. Q., & Palaniappan, R. (2016). Multiresolution analysis over graphs for a motor imagery based online BCI game. *Computers in Biology and Medicine, 68*, 21-26. [DOI:10.1016/j.compbiomed.2015.10.016] [PMID]

Athif, M., & Ren, H. (2019). WaveCSP: A robust motor imagery classifier for consumer EEG devices. *Australasian Physical & Engineering Sciences in Medicine, 42*(1), 159-168. [DOI:10.1007/s13246-019-00721-0] [PMID]

Beisteiner, R., Höllinger, P., Lindinger, G., Lang, W., & Berthoz, A. (1995). Mental representations of movements. Brain potentials associated with imagination of hand movements. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section, 96*(2), 183-193. [DOI:10.1016/0168-5597(94)00226-5] [PMID]

Billinger, M., Brunner, C., & Müller-Putz, G. R. (2013). Single-trial connectivity estimation for classification of motor imagery data. *Journal of Neural Engineering, 10*(4), 046006. [DOI:10.1088/1741-2560/10/4/046006] [PMID]

Brunner, C., Scherer, R., Graimann, B., Supp, G., & Pfurtscheller, G. (2006). Online control of a brain-computer interface using phase synchronization. *IEEE Transactions on Biomedical Engineering, 53*(12), 2501-2506. [DOI:10.1109/TBME.2006.881775] [PMID]

Chen, C., Zhang, J., Belkacem, A. N., Zhang, S., Xu, R., & Hao, B., et al. (2019). G-causality brain connectivity differences of finger movements between motor execution and motor imagery. *Journal of Healthcare Engineering, 2019*, 5068283. [DOI:10.1155/2019/5068283] [PMID] [PMCID]

Cincotti, F., Mattia, D., Aloise, F., Bufalari, S., Schalk, G., & Oriolo, G., et al. (2008). Non-invasive brain-computer interface system: towards its application as assistive technology. *Brain Research Bulletin, 75*(6), 796-803. [DOI:10.1016/j.brainresbull.2008.01.007] [PMID] [PMCID]

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods, 134*(1), 9-21. [DOI:10.1016/j.jneumeth.2003.10.009] [PMID]

Erdoğan, S. B., Yücel, M. A., & Akın, A. (2014). Analysis of task-evoked systemic interference in fNIRS measurements: Insights from fMRI. *Neuroimage, 87*, 490-504. [DOI:10.1016/j.neuroimage.2013.10.024] [PMID]

Friston K. J. (2011). Functional and effective connectivity: A review. *Brain Connectivity, 1*(1), 13–36. [DOI:10.1089/brain.2011.0008] [PMID]

Friston, K. J., Harrison, L., & Penny, W. (2003). Dynamic causal modelling. *Neuroimage, 19*(4), 1273-1302. [DOI:10.1016/S1053-8119(03)00202-7] [PMID]

Gagnon, L., Yücel, M. A., Boas, D. A., & Cooper, R. J. (2014). Further improvement in reducing superficial contamination in NIRS using double short separation measurements. *Neuroimage, 85*, 127-135. [DOI:10.1016/j.neuroimage.2013.01.073] [PMID] [PMCID]

Ginter Jr, J., Blinowska, K. J, Kamiński, M., & Durka, P. (2001). Phase and amplitude analysis in time-frequency space-application to voluntary finger movement. *Journal of Neuroscience Methods, 110*(1-2), 113-124. [DOI:10.1016/S0165-0270(01)00424-1] [PMID]

Ginter Jr, J., Blinowska, K. J., Kamiński, M., Durka, P. J., Pfurtscheller, G., & Neuper, C. (2005). Propagation of EEG activity in the beta and gamma band during movement imagery in humans. *Methods of Information in Medicine, 44*(01), 106-113. [DOI:10.1055/s-0038-1633932]

Gonuguntla, V., Wang, Y., & Veluvolu, K. C. (2016). Event-related functional network identification: Application to EEG classification. *IEEE Journal of Selected Topics in Signal Processing, 10*(7), 1284-1294. [DOI:10.1109/JSTSP.2016.2602007]

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica, 37*(3), 424-438. [DOI:10.2307/1912791]

Granger, C. W. J. (1988). Some recent developments in a concept of causality. *Journal of Econometrics, 39*(1-2), 199-211. [DOI:10.1016/0304-4076(88)90045-0]

Gu, L., Yu, Z., Ma, T., Wang, H., Li, Z., & Fan, H. (2020). EEG-based classification of lower limb motor imagery with brain network analysis. *Neuroscience, 436*, 93–109. [DOI:10.1016/j.neuroscience.2020.04.006] [PMID]

Gu, Q., Li, Z., & Han, J. (2012). Generalized fisher score for feature selection. Retrived from: [Link]

Hamedi, M., Salleh, S. H., Samdin, S. B., & Noor, A. M. (2015). *Motor imagery brain functional connectivity analysis via coherence*. Paper presented at: 2015 IEEE International conference on signal and image processing applications, Kuala Lumpur, Malaysia, 21 October 2015.[DOI:10.1109/ICSIPA.2015.7412202]

Friston, K. J., Harrison, L., & Penny, W. (2003). Dynamic causal modelling. *NeuroImage, 19*(4), 1273–1302. [DOI:10.1016/S1053-8119(03)00202-7] [PMID]

He, L., Hu, D., Wan, M., Wen, Y., Von Deneen, K. M., & Zhou, M. (2015). Common bayesian network for classification of EEG-based multiclass motor imagery BCI. *IEEE Transactions on Systems, Man, and Cybernetics, 46*(6), 843-854. [DOI:10.1109/TSMC.2015.2450680]

He, X., Cai, D., & Niyogi, P. (2005). Laplacian score for feature selection. *Advances in Neural Information Processing Systems, 18*, 1-8. [Link]

Hermes, D., Vansteensel, M. J., Albers, A. M., Bleichner, M. G., Benedictus, M. R., & Mendez Orellana, C., et al. (2011). Functional MRI-based identification of brain areas involved in motor imagery for implantable brain-computer interfaces. *Journal of Neural Engineering, 8*(2), 025007. [DOI:10.1088/1741-2560/8/2/025007] [PMID]

Hsu W. Y. (2015). Enhancing the performance of motor imagery EEG classification using phase features. *Clinical EEG and Neuroscience, 46*(2), 113–118. [DOI:10.1177/1550059413514974] [PMID]

Jansen, B. H., Bourne, J. R., & Ward, J. W. (1981). Autoregressive estimation of short segment spectra for computerized EEG analysis. *IEEE Transactions on Bio-Medical Engineering, 28*(9), 630–638. [DOI:10.1109/TBME.1981.324753] [PMID]

Kamiński, M., Ding, M., Truccolo, W. A., & Bressler, S. L. (2001). Evaluating causal relations in neural systems: Granger causality, directed transfer function and statistical assessment of significance. *Biological Cybernetics, 85*(2), 145–157. [DOI:10.1007/s004220000235] [PMID]

Kamiński, M. J., & Blinowska, K. J. (1991). A new method of the description of the information flow in the brain structures. *Biological Cybernetics, 65*(3), 203–210. [DOI:10.1007/BF00198091] [PMID]

Kim, C., Sun, J., Liu, D., Wang, Q., & Paek, S. (2018). An effective feature extraction method by power spectral density of EEG signal for 2-class motor imagery-based BCI. *Medical & Biological Engineering & Computing, 56*(9), 1645–1658.[DOI:10.1007/s11517-017-1761-4] [PMID]

Kuś, R., Kamiński, M., & Blinowska, K. J. (2004). Determination of EEG activity propagation: Pair-wise versus multichannel estimate. *IEEE Transactions on Bio-Medical Engineering, 51*(9), 1501–1510. [DOI:10.1109/TBME.2004.827929] [PMID]

Liang, S., Choi, K. S., Qin, J., Wang, Q., Pang, W. M., & Heng, P. A. (2016). Discrimination of motor imagery tasks via information flow pattern of brain connectivity. *Technology and Health Care, 24* (Suppl 2), S795–S801. [DOI:10.3233/THC-161212] [PMID]

Liu, H., & Motoda, H. (2008). *Computational methods of feature selection*. Boca Raton: CRC Press. [Link]

Maghsoudi, A., & Shalbaf, A. (2022). Hand motor imagery classification using effective connectivity and hierarchical machine learning in EEG signals. *Journal of Biomedical Physics & Engineering, 12*(2), 161–170. [DOI:10.31661/jbpe.v0i0.1264] [PMID] [PMCID]

Matthews, F., Pearlmutter, B. A., Wards, T. E., Soraghan, C., & Markham, C. (2007). Hemodynamics for brain-computer interfaces. *IEEE Signal Processing Magazine, 25*(1), 87-94. [DOI:10.1109/MSP.2008.4408445]

Mellinger, J., Schalk, G., Braun, C., Preissl, H., Rosenstiel, W., & Birbaumer, N., et al. (2007). An MEG-based brain-computer interface (BCI). *NeuroImage, 36*(3), 581–593.[DOI:10.1016/j.neuroimage.2007.03.019] [PMID] [PMCID]

Meyer, D., Leisch, F., & Hornik, K. (2003). The support vector machine under test. *Neurocomputing, 55*(1-2), 169-186. [DOI:10.1016/S0925-2312(03)00431-4]

Nalatore, H., Ding, M., & Rangarajan, G. (2007). Mitigating the effects of measurement noise on Granger causality. *Physical Review. E, Statistical, Nonlinear, and Soft Matter Physics, 75*(3 Pt 1), 031123. [DOI:10.1103/PhysRevE.75.031123] [PMID]

Naseer, N., & Hong, K. S. (2015). Corrigendum "fNIRS-based brain-computer interfaces: A review". *Frontiers in Human Neuroscience, 9*, 172. [DOI:10.3389/fnhum.2015.00172] [PMID] [PMCID]

Nolte, G., Ziehe, A., Nikulin, V. V., Schlögl, A., Krämer, N., & Brismar, T., et al. (2008). Robustly estimating the flow direction of information in complex physical systems. *Physical Review Letters, 100*(23), 234101. [DOI:10.1103/PhysRevLett.100.234101] [PMID]

Park, S. H., Lee, D., & Lee, S. G. (2018). Filter bank regularized common spatial pattern ensemble for small sample motor imagery classification. *IEEE Transactions on Neural Systems and Rehabilitation Engineering, 26*(2), 498–505. [DOI:10.1109/TNSRE.2017.2757519] [PMID]

Qin, Y., Xu, P., & Yao, D. (2010). A comparative study of different references for EEG default mode network: The use of the infinity reference. *Clinical Neurophysiology, 121*(12), 1981–1991. [DOI:10.1016/j.clinph.2010.03.056] [PMID]

Rathee, D., Cecotti, H., & Prasad, G. (2016). *Estimation of effective fronto-parietal connectivity during motor imagery using partial granger causality analysis*. Paper presented at: 2016 International Joint Conference on Neural Networks., Vancouver, BC, Canada, 29 July 2016. [DOI:10.1109/IJCNN.2016.7727452]

Rehan, M., & Hong, K. S. (2012). Robust synchronization of delayed chaotic FitzHugh-Nagumo neurons under external electrical stimulation. *Computational and Mathematical Methods in Medicine, 2012*, 230980. [DOI:10.1155/2012/230980] [PMID] [PMCID]

Roffo, G. (2016). Feature selection library (MATLAB toolbox). Retrived from: [Link]

Ruan, J., Wu, X., Zhou, B., Guo, X., & Lv, Z. (2018). An automatic channel selection approach for ICA-based motor imagery brain computer interface. *Journal of Medical Systems, 42*(12), 253. [DOI:10.1007/s10916-018-1106-3] [PMID]

Rubinov, M., & Sporns, O. (2010). Complex network measures of brain connectivity: Uses and interpretations. *NeuroImage, 52*(3), 1059–1069. [DOI:10.1016/j.neuroimage.2009.10.003] [PMID]

Sameshima, K., & Baccalá, L. A. (1999). Using partial directed coherence to describe neuronal ensemble interactions. *Journal of Neuroscience Methods, 94*(1), 93–103. [DOI:10.1016/S0165-0270(99)00128-4] [PMID]

Santamaria, L., & James, C. (2016). *Use of graph metrics to classify motor imagery based BCI*. Paper presented at: 2016 International Conference for Students on Applied Engineering, Newcastle Upon Tyne, UK, 21 October 2016. [DOI:10.1109/ICSAE.2016.7810237]

Schreiber T. (2000). Measuring information transfer. *Physical Review Letters, 85*(2), 461–464. [DOI:10.1103/PhysRevLett.85.461] [PMID]

Sharma, N., Pomeroy, V. M., & Baron, J. C. (2006). Motor imagery: A backdoor to the motor system after stroke? *Stroke, 37*(7), 1941–1952. [DOI:10.1161/01.STR.0000226902.43357.fc] [PMID]

Shin, J., von Lühmann, A., Blankertz, B., Kim, D. W., Jeong, J., & Hwang, H. J., et al. (2016). Open access dataset for EEG+NIRS single-trial classification. *IEEE Transactions on Neural Systems and Rehabilitation Engineering, 25*(10), 1735-1745. [DOI:10.1109/TNSRE.2016.2628057]

Shin, Y., Lee, S., Lee, J., & Lee, H. N. (2012). Sparse representation-based classification scheme for motor imagery-based brain-computer interface systems. *Journal of Neural Engineering, 9*(5), 056002. [DOI:10.1088/1741-2560/9/5/056002] [PMID]

Sitaram, R., Weiskopf, N., Caria, A., Veit, R., Erb, M., & Birbaumer, N. (2007). fMRI brain-computer interfaces. *IEEE Signal Processing Magazine, 25*(1), 95-106. [DOI:10.1109/MSP.2008.4408446]

Spiegler, A., Graimann, B., & Pfurtscheller, G. (2004). Phase coupling between different motor areas during tongue-movement imagery. *Neuroscience Letters, 369*(1), 50–54. [DOI:10.1016/j.neulet.2004.07.054] [PMID]

Sun, H., Fu, Y., Xiong, X., Yang, J., Liu, C., & Yu, Z. (2015). [Identification of EEG induced by motor imagery based on hilbert-huang transform (Chinese)]. *Acta Automatica Sinica, 41*(9), 1686-1692. [DOI:10.16383/j.aas.2015.cl50007]

Tharwat, A., Gaber, T., Ibrahim, A. & Hassanien, A.E. (2017). Linear discriminant analysis: A detailed tutorial. *AI Communications*, 30(2), 169-190. [DOI:10.3233/AIC-170729]

Vicente, R., Wibral, M., Lindner, M., & Pipa, G. (2011). Transfer entropy--a model-free measure of effective connectivity for the neurosciences. *Journal of Computational Neuroscience, 30*(1), 45–67. [DOI:10.1007/s10827-010-0262-3] [PMID] [PMCID]

Vickers N. J. (2017). Animal communication: When i'm calling you, will you answer too? *Current Biology, 27*(14), R713–R715. [DOI:10.1016/j.cub.2017.05.064] [PMID]

Vinck, M., Oostenveld, R., van Wingerden, M., Battaglia, F., & Pennartz, C. M. (2011). An improved index of phase-synchronization for electrophysiological data in the presence of volume-conduction, noise and sample-size bias. *NeuroImage, 55*(4), 1548–1565. [DOI:10.1016/j.neuroimage.2011.01.055] [PMID]

Weinrich, M., & Wise, S. P. (1982). The premotor cortex of the monkey. *The Journal of Neuroscience, 2*(9), 1329–1345. [DOI:10.1523/JNEUROSCI.02-09-01329.1982] [PMID] [PMCID]

Wolpaw, J. R., Birbaumer, N., Heetderks, W. J., McFarland, D. J., Peckham, P. H., & Schalk, G., et al. (2000). Brain-computer interface technology: A review of the first international meeting. *IEEE Transactions on Rehabilitation Engineering, 8*(2), 164–173. [DOI:10.1109/TRE.2000.847807] [PMID]

Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., & Vaughan, T. M. (2002). Brain-computer interfaces for communication and control. *Clinical Neurophysiology, 113*(6), 767–791. [DOI:10.1016/S1388-2457(02)00057-3] [PMID]

Xu, L., Zhang, H., Hui, M., Long, Z., Jin, Z., & Liu, Y., et al. (2014). Motor execution and motor imagery: A comparison of functional connectivity patterns based on graph theory. *Neuroscience, 261*, 184–194. [DOI:10.1016/j.neuroscience.2013.12.005] [PMID]

Yavuz, E., & Aydemir, Ö. (2017). *Classification of EEG based BCI signals imagined hand closing and opening*. Paper presented at: The 2017 40th International Conference on Telecommunications and Signal Processing, Barcelona, Spain, 7 July 2017. [DOI:10.1109/TSP.2017.8076020]

Yin, X., Xu, B., Jiang, C., Fu, Y., Wang, Z., & Li, H., et al. (2015). A hybrid BCI based on EEG and fNIRS signals improves the performance of decoding motor imagery of both force and speed of hand clenching. *Journal of Neural Engineering, 12*(3), 036004. [DOI:10.1088/1741-2560/12/3/036004] [PMID]

Zhang, Z., Duan, F., Sole-Casals, J., Dinares-Ferran, J., Cichocki, A., & Yang, Z., et al. (2019). A novel deep learning approach with data augmentation to classify motor imagery signals. *IEEE Access, 7*, 15945-15954. [DOI:10.1109/ACCESS.2019.2895133]