Accepted Manuscript

Accepted Manuscript (Uncorrected Proof)

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

Authors: Erfan Rezaei¹, Ahmad Shalbaf^{1*}

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

***Corresponding author:** Ahmad Shalbaf, Department of Medical Physics and Biomedical Engineering, School of Medicine, Shahid Beheshti University of Medical Sciences, Tehran, Iran. Email: shalbaf@sbmu.ac.ir

To appear in: *Basic and Clinical Neuroscience*

Received date: 2021/02/8

Revised date: 2021/05/29

Accepted date: 2021/09/18

This is a “Just Accepted” manuscript, which has been examined by the peer-review process and has been accepted for publication. A “Just Accepted” manuscript is published online shortly after its acceptance, which is prior to technical editing and formatting and author proofing. *Basic and Clinical Neuroscience* provides “Just Accepted” as an optional and free service which allows authors to make their results available to the research community as soon as possible after acceptance. After a manuscript has been technically edited and formatted, it will be removed from the “Just Accepted” Web site and published as a published article. Please note that technical editing may introduce minor changes to the manuscript text and/or graphics which may affect the content, and all legal disclaimers that apply to the journal pertain.

Please cite this article as:

Rezaei, E., & Shalbaf, A. (In Press). Classification of Right/Left Hand Motor Imagery by Effective Connectivity Based on Transfer Entropy in EEG Signal. *Basic and Clinical Neuroscience*. Just Accepted publication Oct. 2, 2021. Doi: <http://dx.doi.org/10.32598/bcn.2021.2034.3>

DOI: <http://dx.doi.org/10.32598/bcn.2021.2034.3>

Abstract

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 left hand MI task. 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 learning based clustering (LLCFS) algorithms are used to choose the most significant effective connectivity features and reduce redundant information. Finally, support vector machine (SVM) and LDA methods are used for classification. Results show that the best performance in 29 healthy subjects and 60 trials is achieved using the TE method via Relief-F algorithm as feature selection and SVM classification with 91.02% accuracy. Consequently, TE index and a hierarchical feature selection and classification could be useful for discrimination of right and left hand MI task from multichannel EEG signal.

Keywords: Electroencephalogram (EEG), Motor Imagery, Effective Connectivity, Transfer Entropy.

1. Introduction

Brain-computer interface (BCI) is a method (1, 2) that helps paralyzed people to communicate with the environment without using the peripheral nervous system and muscles (3). Electroencephalography (EEG)(4), magnetoencephalography (MEG)(5), fMRI (6), and near-infrared spectroscopy (NIRS)(7) are noninvasive 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 because of being noninvasive, low cost, and high temporal resolution and widely used in neuroscience applications (8-10). 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 paralyzed people for rehabilitation (11). 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 for MI classification. Power spectral density (12), discrete wavelet transform (13), autoregressive model coefficients(14), common spatial pattern (CSP) (15, 16), sparse representation (17), and Hilbert transform (18) were employed for feature extraction from EEG signals. Despite significant results, none of these methods have been proved to be adequately reliable in the practical settings because features from single-channel EEG during MI task can not attain reliable information and the features based on multi-channel EEG 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 (21). PLV was calculated for event related desynchronization/synchronization(ERD/S) between two types of MI tasks (22) and for the phase coupling of sensorimotor during tongue-MI task (23). Santamaria and James employed PLV and wavelet coherence for classifying two different MI tasks with six different classification algorithms (24). PLV were also used directly to categorize left and right hand MI (25) and during left and right hand, foot, and tongue imagery (26). Finally, Hamed et al. used coherence for classifying of four distinct MI tasks (27).

Another popular brain connectivity that is widely used in neuroscience is effective connectivity which refers to the influence that one neural system exerts over another (28). 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) (29), dynamic causal modeling (DCM)(30), and Granger Causality (GC)(31). However, SEM is not appropriate for time series, in which the time constants of neuronal hemodynamics are generally much larger than the fluctuating inputs that drive them (28). 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 are required (32). 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 (33-36). PDC and DTF are able to determine the direction and spectral characteristics of the EEG signals simultaneously (37, 38). Liang et al. employed PDC combined with multivariate empirical mode decomposition during left/right hand MI task to improve classification performance (36). DTF has been used to investigate brain activity dynamics (39) and evaluate motor task experiments (40). The pattern of EEG in beta and gamma band during left and right hand movement imagination was also investigated by DTF (41, 42). Generalized partial directed coherence (GPDC), partial directed coherence factor (PDCF), and full-frequency DTF(ffDTF) are the others extensions of the effective connectivity methods that have been used in the BCI (43).

In summary, these effective connectivity methods have these problems: need a priori information or model, unable to detect nonlinear connections, unable to detect all connectivity in the complex network, and lack of robustness against linear cross-talk between electrophysiological signals (44, 45). So, characterizing and understanding brain dynamics during MI tasks should be done in a way that does not have the mentioned problems. So, an essential nonlinear criterion for estimating effective connectivity with the name of Transfer Entropy (TE) is presented (46). In this method, 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 for analyzing multi-channel EEG signals (47).

After the EEG feature extraction, these features must be optimized by a wide variety of feature selection or dimensional reduction such as principal component analysis (48), independent component analysis (49) and sequential floating forward search (50). Finally, these optimized features must be classified for developing MI tasks such as Bayesian classifier(51), extreme learning machine (52) and deep learning approaches (53). Despite the different machine learning algorithms, there hasn't been a universally superior one 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 task in 29 participants.

2. Material and methods

Participants and experimental design

Twenty-nine healthy subjects (fourteen males and fifteen females with an average of 28.5 ± 3.7 years) with no reported brain-related diseases participated in this study (16). The subjects sat on a comfortable chair at a 1.6-meter distance from 50-inches white screen, and they should not move their body during the task. The experiment consisted of three sessions of right and left-hand MI. Each session consisted of 60 seconds rest before the experiment, twenty repetitions of the task (10 trials for each left and right hand MI), and 60 seconds post-experiment resting period. After the task started with a visual introduction, a task period with ten seconds, and a resting time between fifteen to seventeen seconds, which randomly given is done. In the task period, subjects imagine opening and closing their hands at a speed of 1 Hz. Fig. 1 demonstrates 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 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). Finally, dataset was recorded in Technical University of Berlin and are conducted according to the declaration of Helsinki and approved by the Ethics Committee of the Institute of

Psychology and Ergonomics, Technical University of Berlin (approval number: SH_01_20150330). The data is public in following address: <http://doc.ml.tu-berlin.de/hBCI>.

Preprocessing

All data pre-processing was done using EEGLab in MATLAB R2018b. The measured EEG data filtered by 1Hz FIR high pass filter and re-referenced using a common average reference. ICA based EOG rejection was performed by toolbox in EEGLAB (54).

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 (28). In this study, TE was used to estimate nonlinear effective connectivity. All effective connectivity calculation was done in MATLAB (The Mathworks, Inc., Natick, MA, USA) via the open-source HERMES toolbox.

Transfer Entropy (TE)

TE is a nonparametric method for estimating effective connectivity, that can measure all linear and nonlinear causal relationships (55). In this method, the Wiener causality concept and the conditional mutual information are combined. $MI(x, y)$ Shows the Mutual Information and is defined by equation 1.

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

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 (55).

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

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 (55):

Conditional mutual information $MI(x, y|z)$ is dependent on observing the random variable z and is calculated by equation 3 and 4 (55).

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

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

By combining the Wiener causality and the $MI(x, y|z)$, TE is obtained (55). $TE(x \rightarrow y)$ or TE_{xy} expressions that by assumption knowing the past statement of the random variable x , how much it adds to the available information about the random variable y (55).

$$5) \quad TE(x \rightarrow y) = TE_{xy} = MI\left(y(t + \tau), \underline{x}_t^{d_x, \tau_x} \middle| \underline{y}_t^{d_y, \tau_y}\right)$$

$$\underline{x}_t^{d_x, \tau_x} = (x(t), x(t - \tau_x), \dots, x(t - (d_x - 1)\tau_x))$$

$$\underline{y}_t^{d_y, \tau_y} = (y(t), y(t - \tau_y), \dots, y(t - (d_y - 1)\tau_y))$$

$(\underline{x}_t^{d_x, \tau_x})$ and $(\underline{y}_t^{d_y, \tau_y})$ 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 dimension of x and y , respectively. When $TE=0$, there is no causality 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. These feature reduction and select 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 general characteristics of data, and any predictor and classifiers are ignored (56). 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 (LLCFS) algorithms are employed to choose the best features. Relief-F is an iterative, randomized, and supervised method (57). 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 (58). In the Laplacian score, an unsupervised method, its power of locality preservation evaluates the importance of a feature based on the nearest neighbor graph (59). Finally, the LLCFS algorithm tries to ensure that the cluster tag of each data point is near to the tag predicted by the local regression model, with its adjacent points and their cluster tags (60).

Classification

Support vector machine (SVM) (61) and linear discriminant analysis (LDA) (62) as supervised learning algorithms are used for classification in this study. LDA method is used to find a linear combination of features. SVM as a binary classifier is used to categorize the data so that the margin between the hyperplane and nearest data becomes maximum. In this method, when overlapped features exist, support vector classification maps the feature into a higher dimension space by nonlinear function and create an excellent discriminatory hyperplane in that space. All analyses were computed in MATLAB (The Mathworks, Inc., Natick, MA, USA).

Statistical analysis

10-fold cross validation was used in this study due to the limited dataset. In this method, data divided into 10 parts of equal sizes, and in each run, the classification parameters are constructed with 90 percent of data (80% train and 10% validation (for selecting the optimized number of features)) and tested with 10 percent of data. So, in the first step, we have used only 90 percent of data and 10% of the test data is set aside in each run. When the optimal number of features is selected based on 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. Result

We calculated the effective and functional connectivity between all EEG signals using TE, Coherence and Granger Causality index in 10-sec 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 $30 \times 29 = 870$ trial runs for each class in the classification procedure. We used the Hermes toolbox to extract features. Having 30-channel EEG, 900 (30×30) connection 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 discrimination of left and right MI task. 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). The diagram of the proposed method is shown in Fig 2.

The testing classification accuracies obtained by the TE, Coherence and GC measure and four feature selection methods and classification methods over all participants are given in Table 1, separately. As it can be seen, the proposed method by TE revealed better results rather than other connectivity methods (Coherence and GC). Also, as it 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 Relief-F method yield the best results with high testing classification accuracy (91.02%). It is noteworthy that in TE method, best testing classification accuracy was obtained with smaller number of features rather than GC and Coherence methods. Afterward, TE with SVM and feature selection via the fisher method yield an testing classification accuracy of 86.93. Raw 900 (30*30) connectivity features for the TE method over all left vs. right MI tasks are shown in Figs. 3. In this figure, a higher absolute value of the connectivity feature is shown in warmer colors. We have also plotted P-value of all connectivity features by TE method between left and right hand to show the map of separability in Figure 4. P-values for $30 \times 30 = 900$ features (except 30 diagonal channels) between two classes are calculated and plotted. A lower value of P-value, shown in blue, has more separability. As it can be seen, during the MI task, EEG signals have high separability around the motor area, parietal, and temporal. Fig. 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 it 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 of 12 and 13 features have high average validation classification accuracy for 10 folds. When the optimal number of features is selected based on 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 a 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 task. The novelties of our paper are using the TE method for quantifying the connectivity of the EEG signals during 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.

For estimating the brain effective 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 generally much larger than the fluctuating or exogenous inputs that drive them. GC methods limit the effective pattern to specific templates that are based on the linear parameters of the MVAR model. In comparison, brain connections' natural dynamic is not easy to simply isolate with a predetermined limiting model. Finally, DCM, which partially covers nonlinear interactions, always requires a priori information about the network connections, and this information always does not 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 detect directional connectivity properly.

Motor imagery and motor execution activate the same brain area (63), 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 (47). As shown in Figure 3, 4, differential patterns of effective connectivity between the left vs. right MI task in TE method are around the motor areas rather than other areas. These discriminative features are as well as previous studies (64). 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 (65).

In this study, four widely feature 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. 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, it searches for K nearest neighbors 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 our work results with the prior studies that employed the same data set in EEG signal for classification of left vs. right hand MI task (16, 66, 67). 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 of 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 a multi-model system's performance based on EEG and near-infrared spectroscopy using HbO and Hb as the hemodynamic responses (68, 69) compared with a single modality might improve the accuracy of hand MI task discrimination.

5. Conclusions

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 for

discrimination of left vs. right MI task from EEG signals, testing accuracy of 91.02% on the 29 participants is achieved.

Acknowledgments

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

Accepted Manuscript (Uncorrected Proof)

References

- 1) Yin X, Xu B, Jiang C, Fu Y, Wang Z, Li H, et al. 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*. 2015;12(3):036004.
- 2) Matthews F, Pearlmutter BA, Wards TE, Soraghan C, Markham C. Hemodynamics for brain-computer interfaces. *IEEE Signal Processing Magazine*. 2007;25(1):87-94.
- 3) Wolpaw JR, Birbaumer N, McFarland DJ, Pfurtscheller G, Vaughan TM. Brain-computer interfaces for communication and control. *Clinical neurophysiology*. 2002;113(6):767-91.
- 4) Wolpaw JR, Birbaumer N, Heetderks WJ, McFarland DJ, Peckham PH, Schalk G, et al. Brain-computer interface technology: a review of the first international meeting. *IEEE transactions on rehabilitation engineering*. 2000;8(2):164-73.
- 5) Mellinger J, Schalk G, Braun C, Preissl H, Rosenstiel W, Birbaumer N, et al. An MEG-based brain-computer interface (BCI). *Neuroimage*. 2007;36(3):581-93.
- 6) Sitaram R, Weiskopf N, Caria A, Veit R, Erb M, Birbaumer N. fMRI brain-computer interfaces. *IEEE Signal processing magazine*. 2007;25(1):95-106.
- 7) Naseer N, Hong K-S. fNIRS-based brain-computer interfaces: a review. *Frontiers in human neuroscience*. 2015;9:3.
- 8) A Saeedi, M Saeedi, A Maghsoudi, A Shalbaf, Major depressive disorder diagnosis based on effective connectivity in EEG signals: A convolutional neural network and long short-term memory approach. *Cognitive Neurodynamics* 15 (2), 239-252
- 9) A Shalbaf, S Bagherzadeh, A Maghsoudi, Transfer learning with deep convolutional neural network for automated detection of schizophrenia from EEG signals. *Physical and Engineering Sciences in Medicine* 43 (4), 1229-1239
- 10) F Afshani, A Shalbaf, R Shalbaf, J Sleight, Frontal-temporal functional connectivity of EEG signal by standardized permutation mutual information during anesthesia. *Cognitive neurodynamics* 13 (6), 531-540
- 11) Sharma N, Pomeroy VM, Baron J-C. Motor imagery: a backdoor to the motor system after stroke? *Stroke*. 2006;37(7):1941-52.
- 12) Kim C, Sun J, Liu D, Wang Q, Paek S. An effective feature extraction method by power spectral density of EEG signal for 2-class motor imagery-based BCI. *Medical & biological engineering & computing*. 2018;56(9):1645-58.
- 13) Athif M, Ren H. WaveCSP: a robust motor imagery classifier for consumer EEG devices. *Australasian physical & engineering sciences in medicine*. 2019;42(1):159-68.
- 14) Jansen BH, Bourne JR, Ward JW. Autoregressive estimation of short segment spectra for computerized EEG analysis. *IEEE Transactions on Biomedical Engineering*. 1981(9):630-8.
- 15) Park S-H, Lee D, Lee S-G. Filter bank regularized common spatial pattern ensemble for small sample motor imagery classification. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*. 2017;26(2):498-505.

- 16) Shin J, von Lühmann A, Blankertz B, Kim D-W, Jeong J, Hwang H-J, et al. Open access dataset for EEG+ NIRS single-trial classification. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*. 2016;25(10):1735-45.
- 17) Shin Y, Lee S, Lee J, Lee H-N. Sparse representation-based classification scheme for motor imagery-based brain-computer interface systems. *Journal of neural engineering*. 2012;9(5):056002.
- 18) Sun H, Fu Y, Xiong X, Yang J, Liu C, Yu Z. Identification of EEG induced by motor imagery based on hilbert-huang transform. *Acta Automatica Sinica*. 2015;41(9):1686-92.
- 19) Park H-J, Friston K. Structural and functional brain networks: from connections to cognition. *Science*. 2013;342(6158).
- 20) Gerstein GL, Perkel DH. Simultaneously recorded trains of action potentials: analysis and functional interpretation. *Science*. 1969;164(3881):828-30.
- 21) Qin Y, Xu P, Yao D. A comparative study of different references for EEG default mode network: the use of the infinity reference. *Clinical neurophysiology*. 2010;121(12):1981-91.
- 22) Gu L, Yu Z, Ma T, Wang H, Li Z, Fan H. EEG-based classification of lower limb motor imagery with brain network analysis. *Neuroscience*. 2020;436:93-109.
- 23) Spiegler A, Graimann B, Pfurtscheller G. Phase coupling between different motor areas during tongue-movement imagery. *Neuroscience letters*. 2004;369(1):50-4.
- 24) Santamaria L, James C, editors. Use of graph metrics to classify motor imagery based BCI. 2016 International Conference for Students on Applied Engineering (ICSAE); 2016: IEEE.
- 25) Gonuguntla V, Wang Y, Veluvolu KC. Event-related functional network identification: application to EEG classification. *IEEE journal of selected topics in signal processing*. 2016;10(7):1284-94.
- 26) Brunner C, Scherer R, Graimann B, Supp G, Pfurtscheller G. Online control of a brain-computer interface using phase synchronization. *IEEE Transactions on Biomedical Engineering*. 2006;53(12):2501-6.
- 27) Hamed M, Salleh S-H, Samdin SB, Noor AM, editors. Motor imagery brain functional connectivity analysis via coherence. 2015 IEEE International conference on signal and image processing applications (ICSIPA); 2015: IEEE.
- 28) Friston KJ. Functional and effective connectivity: a review. *Brain connectivity*. 2011;1(1):13-36.
- 29) Anderson JC. Gerbing. DW (1988). Structural equation modeling in practice: A review and recommended two-step approach *Psychological Bulletin*. 103(3):411-23.
- 30) Harrison L, Penny W. Dynamic causal modelling. *Neuroimage*. 2003;19(4):1273-302.
- 31) Granger CW. Some recent developments in a concept of causality," *Journal of Econometrics*, vol. 39, 1988, pp. 199—211; and. Investigating causal relations by econometric models and cross-spectral methods," *Econometrica*. 1969;37:424-38.
- 32) Friston KJ, Harrison L, Penny W. Dynamic causal modelling. *Neuroimage*. 2003;19(4):1273-302.
- 33) Granger CW. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: journal of the Econometric Society*. 1969:424-38.

- 34) Rathee D, Cecotti H, Prasad G, editors. Estimation of effective fronto-parietal connectivity during motor imagery using partial granger causality analysis. 2016 International Joint Conference on Neural Networks (IJCNN); 2016: IEEE.
- 35) Chen C, Zhang J, Belkacem AN, Zhang S, Xu R, Hao B, et al. G-causality brain connectivity differences of finger movements between motor execution and motor imagery. *Journal of healthcare engineering*. 2019;2019.
- 36) Liang S, Choi K-S, Qin J, Wang Q, Pang W-M, Heng P-A. Discrimination of motor imagery tasks via information flow pattern of brain connectivity. *Technology and Health Care*. 2016;24(s2):S795-S801.
- 37) Sameshima K, Baccalá LA. Using partial directed coherence to describe neuronal ensemble interactions. *Journal of neuroscience methods*. 1999;94(1):93-103.
- 38) Kaminski MJ, Blinowska KJ. A new method of the description of the information flow in the brain structures. *Biological cybernetics*. 1991;65(3):203-10.
- 39) Kamiński M, Ding M, Truccolo WA, Bressler SL. Evaluating causal relations in neural systems: Granger causality, directed transfer function and statistical assessment of significance. *Biological cybernetics*. 2001;85(2):145-57.
- 40) Ginter Jr J, Blinowska K, Kamiński M, Durka P. Phase and amplitude analysis in time–frequency space—application to voluntary finger movement. *Journal of neuroscience methods*. 2001;110(1-2):113-24.
- 41) Ginter Jr J, Blinowska K, Kamiński M, Durka P, Pfurtscheller G, Neuper C. Propagation of EEG activity in the beta and gamma band during movement imagery in humans. *Methods of information in medicine*. 2005;44(01):106-13.
- 42) Kus R, Kaminski M, Blinowska KJ. Determination of EEG activity propagation: pair-wise versus multichannel estimate. *IEEE transactions on Biomedical Engineering*. 2004;51(9):1501-10.
- 43) Billinger M, Brunner C, Müller-Putz GR. Single-trial connectivity estimation for classification of motor imagery data. *Journal of neural engineering*. 2013;10(4):046006.
- 44) Nalatore H, Ding M, Rangarajan G. Mitigating the effects of measurement noise on Granger causality. *Physical Review E*. 2007;75(3):031123.
- 45) Nolte G, Ziehe A, Nikulin VV, Schlögl A, Krämer N, Brismar T, et al. Robustly estimating the flow direction of information in complex physical systems. *Physical review letters*. 2008;100(23):234101.
- 46) Schreiber T. Measuring information transfer. *Physical review letters*. 2000;85(2):461.
- 47) Weinrich M, Wise SP. The premotor cortex of the monkey. *Journal of Neuroscience*. 1982;2(9):1329-45.
- 48) Vickers NJ. Animal communication: when i'm calling you, will you answer too? *Current biology*. 2017;27(14):R713-R5.
- 49) Ruan J, Wu X, Zhou B, Guo X, Lv Z. An automatic channel selection approach for ICA-Based motor imagery brain computer interface. *Journal of Medical Systems*. 2018;42(12):253.
- 50) Asensio-Cubero J, Gan JQ, Palaniappan R. Multiresolution analysis over graphs for a motor imagery based online BCI game. *Computers in biology and medicine*. 2016;68:21-6.

- 51) He L, Hu D, Wan M, Wen Y, Von Deneen KM, Zhou M. Common Bayesian network for classification of EEG-based multiclass motor imagery BCI. *IEEE Transactions on Systems, man, and cybernetics: systems*. 2015;46(6):843-54.
- 52) Hsu W-Y. Enhancing the performance of motor imagery EEG classification using phase features. *Clinical EEG and neuroscience*. 2015;46(2):113-8.
- 53) Zhang Z, Duan F, Sole-Casals J, Dinares-Ferran J, Cichocki A, Yang Z, et al. A novel deep learning approach with data augmentation to classify motor imagery signals. *IEEE Access*. 2019;7:15945-54.
- 54) Delorme A, Makeig S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of neuroscience methods*. 2004;134(1):9-21.
- 55) Vicente R, Wibral M, Lindner M, Pipa G. Transfer entropy—a model-free measure of effective connectivity for the neurosciences. *Journal of computational neuroscience*. 2011;30(1):45-67.
- 56) Roffo G. Feature selection library (MATLAB toolbox). arXiv preprint arXiv:160701327. 2016.
- 57) Liu H, Motoda H. *Computational Methods of Feature Selection* (Chapman & Hall/Crc Data Mining and Knowledge Discovery Series). 2008.
- 58) Gu Q, Li Z, Han J. Generalized fisher score for feature selection. arXiv preprint arXiv:12023725. 2012.
- 59) He X, Cai D, Niyogi P, editors. Laplacian score for feature selection. *Advances in neural information processing systems*; 2006.
- 60) Zeng H, Cheung Y-m. Feature selection and kernel learning for local learning-based clustering. *IEEE transactions on pattern analysis and machine intelligence*. 2010;33(8):1532-47.
- 61) Meyer D, Leisch F, Hornik K. The support vector machine under test. *Neurocomputing*. 2003;55(1-2):169-86.
- 62) Tharwat A, Gaber T, Ibrahim A, Hassanien AE. Linear discriminant analysis: A detailed tutorial. *AI communications*. 2017;30(2):169-90.
- 63) Beisteiner R, Höllinger P, Lindinger G, Lang W, Berthoz A. Mental representations of movements. Brain potentials associated with imagination of hand movements. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*. 1995;96(2):183-93.
- 64) Hermes D, Vansteensel MJ, Albers AM, Bleichner MG, Benedictus MR, Orellana CM, et al. Functional MRI-based identification of brain areas involved in motor imagery for implantable brain–computer interfaces. *Journal of neural engineering*. 2011;8(2):025007.
- 65) Xu L, Zhang H, Hui M, Long Z, Jin Z, Liu Y, et al. Motor execution and motor imagery: a comparison of functional connectivity patterns based on graph theory. *Neuroscience*. 2014;261:184-94.
- 66) Yavuz E, Aydemir Ö, editors. Classification of EEG based BCI signals imagined hand closing and opening. 2017 40th International Conference on Telecommunications and Signal Processing (TSP); 2017: IEEE.

- 67) Maghsoudi A, Shalbaf A. Hand Motor Imagery Classification Using Effective Connectivity and Hierarchical Machine Learning in EEG Signals. *Journal of Biomedical Physics and Engineering*. 2020.
- 68) Erdoğan SB, Yücel MA, Akın A. Analysis of task-evoked systemic interference in fNIRS measurements: insights from fMRI. *NeuroImage*. 2014;87:490-504.
- 69) Gagnon L, Yücel MA, Boas DA, Cooper RJ. Further improvement in reducing superficial contamination in NIRS using double short separation measurements. *Neuroimage*. 2014;85:127-35.

Accepted Manuscript (Uncorrected Proof)

Fig. 1. Schematic diagram of the experimental paradigm.

Fig2. The process of the proposed system. Raw EEG data (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.

Fig 3. Raw 900 (30*30) TE connectivity features over all participants for MI task. A higher absolute value of connectivity feature shows with warm colors. (a) TE EEG features for right hand imagination (b) TE EEG features for left hand imagination.

Fig 4. P-value of all connectivity features by TE method between left and right hand to show the map of separability. A lower value of P-value, shown in blue, has more separability.

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

Table 1: Testing classification accuracy obtained from connectivity methods using TE, Coherence, and Granger Causality over all participants using four feature selection methods (Relief-F, laplacian, LLCFS, fisher) and finally SVM and LDA classification structure. The selected number of features to reach best accuracy on validation dataset in each method was mentioned.

Connectivity method	Classifier	Feature selection methods			
		Relief-F	Laplacian	LLCFS	Fisher
TE	LDA	85.74±0.002 (14 features)	82.31±0.0021 (18 features)	71.69±0.0011 (15 features)	84.9±0.002 (14 features)
	SVM	91.02±0.0015 (12 features)	86.87± 0.0031 (17 features)	85.21±0.0030 (13 features)	86.93 ±0.003 (13 features)
Coherence	LDA	75.12±0.002 (20 features)	66.32±0.002 (25 features)	67.85±0.001 (19 features)	63.45±0.002 (22 features)
	SVM	80.67±0.0021 (17 features)	81.42± 0.003 (23 features)	79.69±0.003 (17 features)	67.42 ±0.003 (23 features)
GC	LDA	80.13±0.0012 (18 features)	82.36±0.0036 (20 features)	74.63±0.0021 (18 features)	73.35±0.002 (17 features)
	SVM	82.48±0.0021 (16 features)	83.64± 0.0026 (18 features)	84.12±0.0029 (17 features)	74.36 ±0.0034 (15 features)

Table 2: Comparison our work results with the prior studies that employed the same data set in EEG signal.

Authors	Dataset	year	feature	classifier	accuracy
Jaeyoung Shin(16)	29subject, Shin dataset	2016	CSP	LDA	65.6
Ebru Yavuz(66)	29subject, Shin dataset	2017	Hilbert Transform	Knn	82.23
				LDA	78.13
Arash Masghsoudi(67)	29subject, Shin dataset	2020	DTF		61.85
			dDTF	SVM	73.66
			GPDC		83.87
Proposed method	29subject, Shin dataset	2021	TE	SVM	91.02
				LDA	85.74