Decoding Articulation Motor Imagery Using Early Connectivity Information in the Motor Cortex: A Functional Near-Infrared Spectroscopy Study

Zengzhi Guo and Fei Chen, *Senior Member, IEEE*

***Abstract*— Brain computer interface (BCI) based on speech imagery can help people with motor disorders communicate their thoughts to the outside world in a natural way. Due to being portable, non-invasive, and safe, functional near-infrared spectroscopy (fNIRS) is preferred for developing BCIs. Previous BCIs based on fNIRS mainly relied on activation information, which ignored the functional connectivity between neural areas. In this study, a 4-class speech imagery BCI based on fNIRS is presented to decode simplified articulation motor imagery (only the movements of jaw and lip were retained) of different vowels. Synchronization information in the motor cortex was extracted as features. In multiclass (four classes) settings, the mean subject-dependent classification accuracies approximated or exceeded 40% in the 0-2.5 s and 0-10 s time windows, respectively. In binary class settings (the average classification accuracies of all pairwise comparisons between two vowels), the mean subject-dependent classification accuracies exceeded 70% in the 0-2.5 s and 0-10 s time windows. These results demonstrate that connectivity features can effectively differentiate different vowels even if the time window size was reduced from 10 s to 2.5 s and the decoding performance in both the time windows was almost the same. This finding suggests that speech imagery BCI based on fNIRS can be further optimized in terms of feature extraction and command generation time reduction. In addition, simplified articulation motor imagery of vowels can be distinguished, and therefore, the potential contribution of articulation motor imagery information extracted from the motor cortex should be emphasized in speech imagery BCI based on fNIRS to improve decoding performance.**

***Index Terms*— Brain computer interface, speech imagery, functional near-infrared spectroscopy, synchronization information, command generation time reduction.**

Manuscript received 18 May 2022; revised 15 September 2022 and 1 December 2022; accepted 4 December 2022. Date of publication 8 December 2022; date of current version 1 February 2023. This work was supported in part by the National Natural Science Foundation of China under Grant 61971212 and in part by the Shenzhen Sustainable Support Program for High-Level University under Grant 20200925154002001. *(Corresponding author: Fei Chen.)*

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Ethical Review Board of Southern University of Science and Technology, Shenzhen, China, under Approval No. 20200068.

Zengzhi Guo is with the School of Electronics and Information Engineering, Harbin Institute of Technology, Harbin 150001, China, and also with the Department of Electrical and Electronic Engineering, Southern University of Science and Technology, Shenzhen 518055, China (e-mail: 11849552@mail.sustech.edu.cn).

Fei Chen is with the Department of Electrical and Electronic Engineering, Southern University of Science and Technology, Shenzhen 518055, China (e-mail: fchen@sustech.edu.cn).

This article has supplementary downloadable material available at https://doi.org/10.1109/TNSRE.2022.3227595, provided by the authors.

Digital Object Identifier 10.1109/TNSRE.2022.3227595

I. Introduction

BRAIN computer interface (BCI) is a communication system that allows individuals with severe motor impairments to control computers or robots, without using the normal pathways of muscle activity and peripheral nerves, but using neural signals produced by brain activity [1]. Individuals with very limited or no voluntary movement can interact with their surroundings without vocal speech or physical gestures by a BCI system. The BCI systems with protocols that need users to attend to stimuli are commonly named as reactive BCIs, such as P300 spellers and steady state visual evoked potential spellers [1]. Reactive BCIs rely on additional displays to generate external stimuli which makes them not convenient to be used. Active BCI needs users to perform several mental tasks. Most of the mental tasks used in active BCI, such as mental arithmetic or motor imagery, have little or no correlation with the actual intended message. Speech imagery is a mental task that imagines speaking without any articulation movements [2]. Speech imagery is an intuitive medium to reflect the thought of users, which overcomes the shortcomings of the above mental tasks.

Magnetoencephalography, functional magnetic resonance imaging (fMRI), electroencephalography (EEG), positron emission tomography (PET), and functional near-infrared spectroscopy (fNIRS) are frequently used functional neuroimaging modalities to record brain signals. Among them, fNIRS and EEG are preferred for developing BCIs because of being portable, non-invasive, safe, and low cost [3]. EEG measures brain activity by measuring neuroelectric signals of neurons from electrodes placed on the scalp. Although EEG offers high temporal resolution (e.g., $>$1000 Hz), it suffers from the low spatial resolution (e.g., 5.0-9.0 cm) which limits EEG to accurately localize associated cortical sources [4]. fNIRS, which measures the concentration changes of oxygenated hemoglobin (HbO) and deoxygenated hemoglobin (HbR) through the intact skull based on near-infrared-range

light, offers a higher spatial resolution (∼3.0 cm) than EEG and a relatively better temporal resolution (∼10 Hz) as compared to fMRI. fNIRS can measure brain activity in complex situations as it is not much susceptible to motion artifacts unlike EEG [5].

fNIRS-based speech imagery BCI has been proven to be feasible many times. Tanino et al. [6] used the fNIRS to discriminate the yes or no answers without voice and gesture for ten healthy subjects. The subjects imagined affirmative or negative answers to a variety of questions, with an average offline classification accuracy of 69%. A similar yes/no paradigm also was tested on eight healthy participants with fNIRS. Hwang et al. [7] extracted multiple types of features in different time windows from fNIRS signals and selected the best feature subset from them. An average offline classification accuracy of about 75% was reported. Sereshkeh et al. [8] used fNIRS to distinguish covertly rehearsed yes or no responses and unconstrained rest in a sample of twelve participants and obtained an average online three-class accuracy rate of 64.1%. The feasibility of fNIRS-based speech imagery BCI also has been tested on a patient with amyotrophic lateral sclerosis. Gallegos-Ayala et al. [9] tested a yes/no paradigm on a patient with amyotrophic lateral sclerosis using fNIRS and the online classification accuracy was 71.7%.

Although fNIRS-based BCIs have been developed for more than ten years, they suffer from a lacking of fNIRS-specific algorithm studies as compared to EEG-based BCIs. Hong et al. counted 74 articles (2010-2019) on fNIRS-based BCIs, and they found that 99% of the articles used temporal features, such as mean, kurtosis, variance, and slope [3]. However, detecting and distinguishing brain activity caused by motor imagery is very difficult because the strength of neural activity caused by motor imagery is reduced by 70% than that of voluntary movements [10]. Therefore, more intelligent algorithms are required to detect the small changes of brain activity associated with motor imagery. Some studies based on EEG have observed that synchronization information is more discriminative than activation amplitude information due to the sensitivity of synchronization changes [11], [12]. Connection features, which can reflect information dissemination and processing, have been proven to be effective in various EEG studies, such as emotion detection and motor imagery classification [12], [13], [14], [15]. Motivated by those studies, functional connectivity features were utilized to decode the articulation motor imagery of different vowels.

Slow hemodynamic response makes the time to generate commands of fNIRS-based BCIs too long. A previous fMRI study measured the temporal delay of hemodynamic response and reported that the hemodynamic response peaked within about 6 s after being evoked by visual stimuli [16]. Due to the relatively slow hemodynamic response, the command generation time of BCI based on fNIRS usually took at least 15 s which makes the information transmission rate relatively low [17]. Therefore, rapid decoding of the commands to reduce delays in fNIRS-based BCIs is becoming a research focus. Zafar and Hong demonstrated that the decoding performance of three metal tasks based on the early information extracted from fNIRS signals in a 0-2.5 s time window could reach 57.5% [18]. However, the classification accuracy based on signal mean and peak value is a bit sacrificed from 65.9% to 57.5% when the time window size was reduced from 7 s to 2.5 s. Using fNIRS signals in a 0-2 s time window, Li et al. found the movements of the right and left hands could be distinguished in a sample of eleven healthy subjects and a classification accuracy rate of 85.5% was reported [19]. A reduced window size (0–2 s) also was used in a BCI based on fNIRS to decode four mental tasks (mental arithmetic, mental counting, mental rotation, and word formation) for ten healthy adults, the classification accuracy rate reached 75.6% [20]. Zafar and Hong used early information of fNIRS signal extracted from 0-2.5 s time window to decode two finger tapping tasks (right-hand thumb and little finger) and achieved an average classification accuracy rate of 74.9% [21]. Those results suggested that the early information of fNIRS signal could provide lofty differentiable information for BCIs.

The purpose of this work was to establish a framework for an intuitive and natural communication BCI that utilized the brain activities in the motor cortex caused by simplified articulation motor imagery to decode different Chinese vowels. Features were extracted from synchronization information of fNIRS signals. To reduce the command generation time, the running temporal window size was reduced from 10 s to 2.5 s for the fNIRS-based BCI in this work.

II. Methods

A. Participants

Nineteen (7 women and 12 men; mean age 20 years, range 18-26 years) healthy individuals participated in this study. The study was approved by the Ethical Review Board of Southern University of Science and Technology in Shenzhen China. All subjects were right-handed and had no history of cardiovascular, psychiatric, respiratory, neurological, drug-related, or alcohol-related diseases. Chinese was the native language of all subjects. Participants were given their informed consent to participate and the procedure was accurately explained prior to the experiment.

B. Instrumentation

To measure fNIRS signals from the motor cortex, an imaging system (Lightnirs; Shimadzu, Japan), which had 16 optical fibers (8 sources, 8 detectors), was equipped. A 2 × 8 fNIRS optode array with an emitter-detector separation of 3 cm was placed over the subject's head. The NIRS signals were sampled at 13.33 Hz. Each fNIRS emitter contained two laser diodes that emitted light at wavelengths of 760 nm and 850 nm, respectively. 22 light channels, constituted by 16 fNIRS optical poles, were arranged as shown in Fig. 1. A three-dimensional digitizer (FASTRAK, Polhemus, Vermont, USA) was used to measure the locations of the optical poles. Positions of the 22 light channels in the Montreal Neurological Institute (MNI) standard brain space were estimated for each subject [22].The MNI coordinates of fNIRS channels and channel cortical projection points are listed in Table SI (in Supplementary Material). The measured brain areas are mainly divided into two regions: supplementary motor area (SMA) and premotor cortex (PMC).

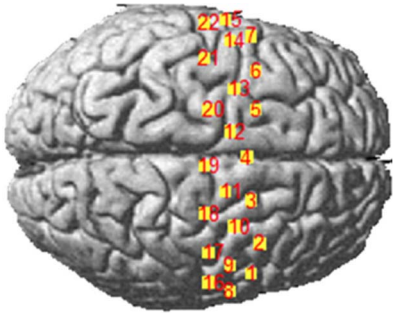

Fig. 1. The positions of light channels.

C. Experimental Protocol

The complicated articulation movements of speaking Chinese vowels were simplified and only movements of the jaw and lips were retained in this study. The movements of the jaw and lips are controlled by different muscles. Jaw opening is controlled by the digastricus, whereas lip rounding is controlled by the orbicularis oris [23]. During the experiment, four stimulus materials were presented as visual cues on a screen and the subjects were instructed by the visual cues to perform the tasks, including: vowel /a/ (imagined jaw opening widely with imagined vocalization); vowel /u/ (imagined lip rounding with imagined vocalization); vowel /i/ (imagined jaw opening slightly with imagined vocalization); and vowel /o/ (imagined jaw opening moderately and lip rounding with imagined vocalization). For vowel /a/, jaw opening widely means opening the mouth to a round shape. For vowel /o/, jaw opening moderately means opening the mouth to an oval shape. For vowel /i/, jaw opening slightly means opening the mouth slightly. Jaw closing means mouth closing.

The subjects were seated on a comfortable chair in a dark soundproof room, about 90 cm away from the display which was used to provide visual instructions. The fNIRS head cover was placed on the subject's head. Before the start of the experiment, the subjects were coached and rehearsed with real movements to ensure correct task execution. 20 trials were performed for imagining movements of each vowel, and therefore a total of 80 imagery trials were performed by each subject. Subjects performed the real movements of each vowel for 20 trials before the imagery tasks to ensure that the subjects could conduct the trials and the imagery tasks correctly. Subjects rested for 5∼10 min after performing 40 trials.

The timing for experimental trials was organized in the following manner as shown in Fig. 2. First, a vowel symbol chosen in a pseudo-random order appeared on the screen, and disappeared after 2 s. Then the vowel symbol was replaced by the symbol #. The symbol # remained on the screen for a time interval of 2 s and was then replaced by the symbol $*$. The $*$ symbol remained on the screen for 10 s and the participants were asked to perform the task repeatedly during this period. When the $*$ symbol was replaced by the # symbol on the screen, the participants had a 15 s rest interval, until the start of the next trial. Participants proceeded to the next trial after the 15 s rest.

D. Signal Pre-Processing

Data quality was evaluated manually by an expert. As the data quality was poor, the data from one subject was excluded from the following signal processing. Signal pre-processing was performed using MATLAB software (Mathworks, Sherborn, MA, USA). First, optical signals were processed using the NIRS-SPM toolbox [24] and converted to HbO and HbR concentration changes by using the modified Beer-Lambert law. Motion artifact segments were replaced with spline interpolation based on neighboring signals to correct motion artifacts [25]. The signals were filtered using a 3rd order Butterworth low-pass filter with a cutoff frequency of 0.4 Hz to remove cardiac interference (0.8 Hz). Long-term baseline drifts were minimized using a 5th order Butterworth high-pass filter with a cutoff frequency of 0.02 Hz.

E. Feature Extraction

In this study, functional connectivity features were extracted for the classification tasks because functional connectivity can reflect information propagation. The temporal correlation of regional hemodynamics was estimated by Pearson's correlation which was commonly used to quantify functional connectivity between different brain regions in fNIRS-based studies [26], [27], [28], [29]. The Pearson correlation coefficient is used to measure the degree of linear correlation between two numerical variables. The Pearson correlation coefficient is obtained by dividing the covariance by the standard deviation of the two variables. The strength of the correlation, as well as the direction of the relationship could be reflected by the Pearson correlation coefficient. The value of Pearson correlation coefficient is assigned between -1 and 1. When the value is 1, it means a total positive correlation; when the value is -1, it means a total negative correlation; when the value is 0, it means no correlation between the two variables. The Pearson correlation coefficient r is calculated as follows:

$$r = \frac{\sum_{i=1}^{n} (X_i - M_x) \cdot (Y_i - M_y)}{\sqrt{\sum_{i=1}^{n} (X_i - M_x)^2} \cdot \sqrt{\sum_{i=1}^{n} (Y_i - M_y)^2}}, \tag{1}$$

where X_i is the i-th point of the HbO/HbR signal of one channel and Y_i is the i-th point of the HbO/HbR signal of another channel. n is the number of data points of the HbO/HbR signal. M_x and M_y are the average values of the HbO/HbR signal for the two channels, respectively.

Correlation coefficients for HbO and HbR concentration time series recorded from each pair of leads were computed. Two 22×22 correlation matrixes were obtained for each trial of each subject. The indices of fNIRS channels were represented by the row and column of the correlation matrices. The correlation coefficients were represented by the matrix elements. Because Pearson's correlation was undirected, the correlation matrix is a symmetric matrix that takes the main diagonal as the axis of symmetry, and all elements are symmetrically equal. Only the elements in the upper triangular

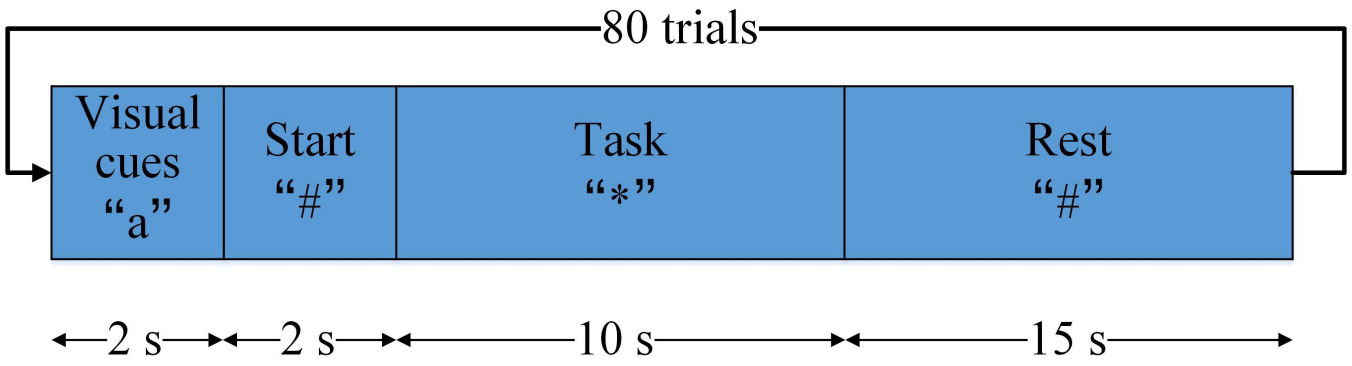

Fig. 2. The timing diagram of the experiment.

part of the correlation matrixes of HbO and HbR signals (except those diagonal elements) were selected as connection features in each trial. Those connection features constituted a one dimensional feature vector. Therefore, the connection features of HbO and HbR signals in each trial were reduced from 968 ($= 2 \times 22 \times 22$) to 462 ($= 2 \times (1 + 2 + \ldots + 21)$).

In order to compare the decoding performance of connection features with that of commonly used activation features, mean, slope, kurtosis, skewness, and variance values of HbO and HbR signals in each channel were obtained as activation features for classification. Mean values were calculated by averaging all data points. Slope values were computed with the polyfit function in Matlab. Kurtosis values were computed with the kurtosis function in Matlab. Skewness values were computed with the skewness function in Matlab. Variance values were computed with the var function in Matlab. However, the number of connection features (462 candidate features) is bigger than that of each type of activation features (44 candidate features).

F. Feature Selection

Feature selection plays an extremely important role in classification problem-solving. On the one hand, in the case of limited samples, it is not appropriate to design a classifier with a large number of features in terms of computational overhead and classifier performance; On the other hand, the relationship between features and classifier performance is not necessarily linear, and when the number of features exceeds a certain limit, the performance of the classifier will deteriorate. Therefore, effective feature selection has become a problem that must be solved in classification. Feature selection can eliminate irrelevant or redundant features, thereby reducing the number of features, improving model accuracy, and reducing running time. The selection of valid features simplifies the model and makes it easier for understanding features.

Fisher score (F-score) is a ranking method to assign each feature with a weight value to indicate the importance degree of the feature. Because the calculation of F-score is simple and fast, it has become one of the most commonly used feature selection methods for BCI based on fNIRS [30], [31]. The impressive classification results of those studies proved that the optimal feature subset selected by F-score had a high distinguishing ability. The main idea of F-score is that the features with strong discriminative power should be as small as possible intra-class distance and as large as possible inter-class distance. The F-score of the i-th feature of the dataset is defined as:

$$FS_i = \frac{\sum_{k=1}^{c} \left(M_{k,i} - M_i\right)^2}{\sum_{k=1}^{c} \frac{1}{(n_k - 1)} \sum_{j=1}^{n_k} \left(x_{k,j,i} - M_i\right)^2}, \tag{2}$$

where M_i, $M_{k,i}$ are the mean values of the i-th feature in the train aset and in the train set of k-th class, respectively, $x_{k,j,i}$ is the i-th feature of the j-th sample in the k-th class, and n_k is the number of samples in k-th class. A high F-score indicates that the feature has a large separability.

Leave-one-out cross-validation (LOOCV) scheme was used to estimate the classification accuracy because it is suitable for small samples and the data set in this study contained only 80 samples. In LOOCV, if there are N samples in the original data set, then each sample is used as a validation set alone, and the remaining N-1 samples are used as a training set, so N classifiers were constructed based on the N training sets; Each validation set was tested with the corresponding classifier; The average classification accuracy of the N classifiers is used as the classification performance indicator. According to the procedure of LOOCV, the sample utilization of it is high because each classifier in it is constructed by almost all samples and therefore the classification accuracy error caused by the inefficient sample is significantly reduced. The classification accuracies of support vector machine (SVM) (Linear, Gaussian, and Polynomial kernels), linear discriminant analysis (LDA), and k-nearest neighbor (KNN) were compared for distinguishing different articulation motor imagery tasks. Those well-known learning algorithms were chosen to classify in this study for their simplicity and low computational requirements, therefore, those algorithms have the potential to be applied in online BCI systems in the future.

A wrapper model with LOOCV scheme and the well-known machine learning algorithms was used to select features in this study. The wrapper model evaluated each feature subset's performance with each machine learning method and selected

the optimal feature subset for each machine learning method with the highest classification accuracy. The entire procedure of the wrapper model for each machine learning method is described as follows:

1. According to the procedure of LOOCV, 80 train sets and the corresponding 80 test sets were obtained.
2. For each train set, each feature in the train set was assigned with an F-score, and all the features were sorted in descending order according to F-score values. N represents the number of features in the rearranged feature set.
3. The first n features with the highest F-score values were selected from the rearranged feature set to establish a new feature subset where n changed from 1 to N. N feature subsets were obtained for each train set.
4. Each machine learning method was used to evaluate the feature subsets with the same number of features in different train sets. The classification accuracies of all feature subsets with different numbers of features were obtained.
5. The best classification result and the corresponding optimal feature subset for each machine learning method were selected.

III. Results

Multiple brain regions are activated by speech imagery and functional connections play a vital role in speech processing. Neural signatures differ between the different imagery tasks in activation and connection patterns as elaborated below. In this study, connection features were used to distinguish different tasks. Multiclass and binary classification accuracy scores were recorded. Furthermore, in order to reduce the time to generate a command, the time window size for extracting features was reduced from 10 s to 2.5 s, and the multiclass and binary classification accuracies were calculated.

A. Neural Signatures in Activation Pattern

In this study, only the neural signatures of HbO signals were shown because there is an extremely negative correlation between HbO signals and HbR signals and the changes of HbR signals can be reflected by that of HbO signals [32]. In order to create the activation map, the 10 s task period data of each trial was extracted. Activation was calculated by subtracting the average value of the HbO in 2 s before the task from the average value of the HbO in 3-7 s during the task. The aim of subtracting the average HbO value in 2 s before the task was to correct baseline. The average value of HbO in 3-7 s during the task was used because the HbO signals usually reached peaks in this period in this study. 22 features were obtained for each trial. The overall average activation level for each type of task was obtained by averaging the activation levels of all trials and all subjects. These activation levels were arranged into spatial maps according to the positions of the channels. Linear interpolation was used to restore the data between different channels. To compare the HbO activation levels of different vowels imagery tasks and that of baseline period, the paired tests were conducted according to Wilcoxon signed-rank test on the HbO activation levels for each channel. This nonparametric test can be applied regardless of the sample distribution and whether the distribution is known. The channels with $p < 0.05$ were considered significantly activated in the imagery tasks.

Figure 3 depicts the activation maps of different vowel imagery tasks. The activation patterns for all tasks are similar and the activation levels of channels 1, 7, 8, 9, 12, 15, 16, and 19 are slightly higher than that of other channels. Channel 8 is significantly activated for vowel /o/, /u/, and /i/ imagery tasks and more channels are significantly activated for vowel /o/ than other vowels. Two large bilateral clusters of high activations were found located on the bilateral PMC. In addition, a small parietal cluster of high activations located on the SMA also was observed. The activation analysis revealed a set of brain areas involved in articulation motor imagery: the bilateral PMC and the SMA. Compared with the other vowels, the activation level for the articulation motor imagery of vowel /o/ is higher. Wilcoxon rank-sum tests were conducted on the activation levels of trials between vowel /o/ and the other vowel. Wilcoxon rank-sum test is a nonparametric test. It does not require the residuals of the fitted model to follow a normal distribution. The channels with significantly higher ($p < 0.05$) activation levels of vowel /o/ than those of the other vowel are marked on the brain maps (Fig. 4). The activation levels of vowel /o/ are significantly higher in some brain areas than that of the other vowels.

B. Neural Signatures in Connection Pattern

Figures 5 and 6 show the average adjacency matrixes of the weighted brain networks across 18 subjects for different tasks in the 0-10 s and 0-2.5 s time windows, respectively. According to the positions of channels on the motor cortex, the channels were divided into three groups. Channels 1, 2, 8, 9, 10, 16, and 17 were located on the left PMC; channels 3, 4, 5, 11, 12, 18, 19, and 20 were located on the SMA; and channels 6, 7, 13, 14, 15, 21, and 22 were located on the right PMC. As presented in Figs. 5 and 6, distributions of the average connection coefficient matrices for the articulation motor imagery of four vowels were almost identical in both 0-10 and 0-2.5 time windows. The brain functional connectivity for motor imagery of four vowels is evenly distributed in both the time windows. The correlations concentrated on the left PMC were stronger than that concentrated on the right PMC. Those results might indicate that the left PMC plays a more important role in articulation motor imagery than the right PMC.

The connection strengths of each vowel were used to subtract those of the other vowels to estimate the connection strength differences. The connection strength differences among paired vowels across all subjects in the 0-10 s and 0-2.5 s time windows are depicted in Figs. 7 and 8, respectively. Wilcoxon rank-sum test was conducted to check whether the connection strengths were significantly different between paired vowels in the different time windows. Figures S1 and S2 (in Supplementary Material) depict the linkages of each vowel with significantly higher connection strengths than those of the other vowels in the 0-10 s and

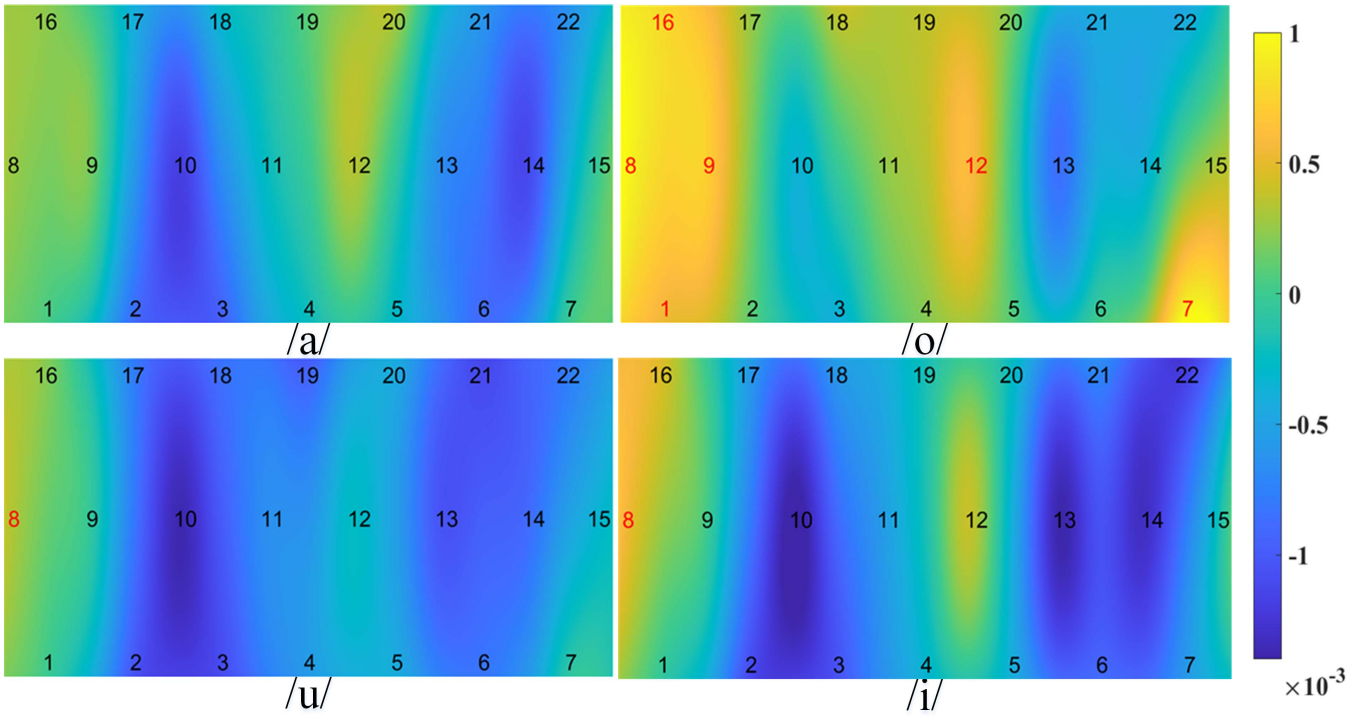

Fig. 3. The brain maps show average HbO activations for articulation motor imagery of different vowels across all subjects. The channel number with the red color indicates that the response of this channel is significantly higher than the baseline ($p < 0.05$, Wilcoxon signed-rank tests).

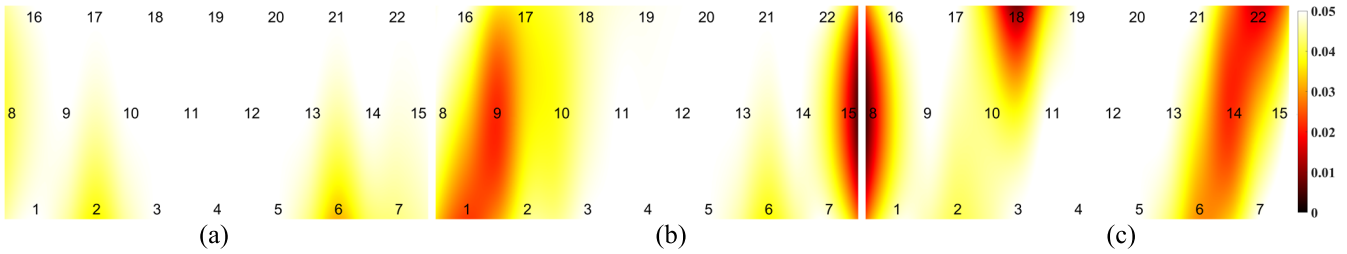

Fig. 4. The statistical mapping for the HbO activations contrast between vowel /o/ and vowel /a/, /u/ or /i/. A warm color indicates that the response in the channel of the vowel /o/ is significantly higher than that of the other vowels ($p < 0.05$, Wilcoxon rank-sum tests). (a) /o/ vs. /a/. (b) /o/ vs. /u/. (c) /o/ vs. /i/.

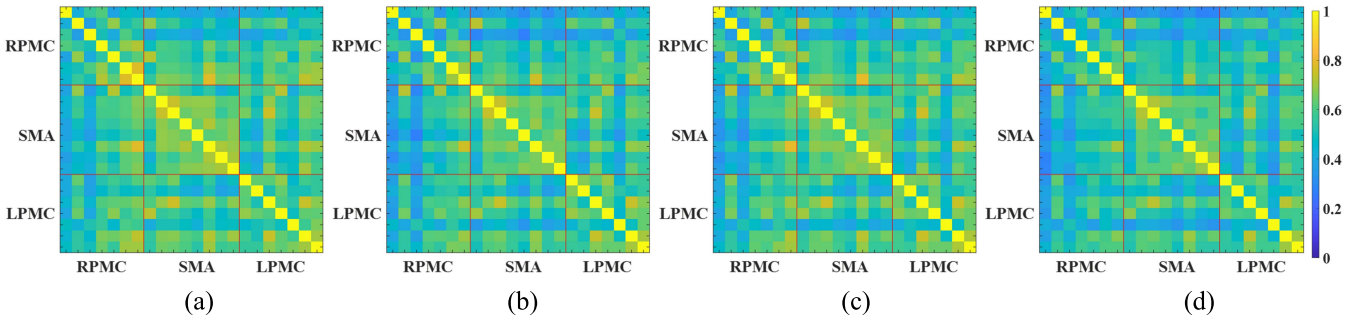

Fig. 5. Average correlation coefficient matrixes across all subjects of HbO in 0-10 s time window. (a) Articulation motor imagery of vowel /a/. (b) Articulation motor imagery of vowel /o/. (c) Articulation motor imagery of vowel /u/. (d) Articulation motor imagery of vowel /i/.

0-2.5 s time windows, respectively. In the 0-10 s time window, most connection strengths of vowel /a/ are higher than those of the other vowels and a lot of connection strengths of vowel /a/ are significantly higher than those of the other vowels. Any connection strength of vowels /o/, /u/, /i/ is not significantly higher than that of vowel /a/. Most connection strengths of vowel /i/ are lower than those of the other vowels and a lot of connection strengths of vowels /a/, /o/, and /u/ are significantly higher than those of vowel /i/. Any connection strength of vowel /i/ is not significantly higher than that of the other

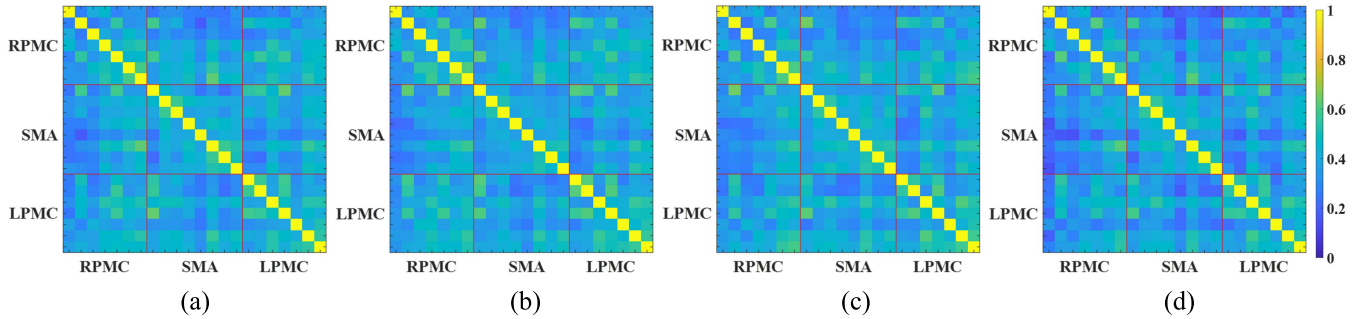

Fig. 6. Average correlation coefficient matrixes across all subjects of HbO in 0-2.5 s time window. (a) Articulation motor imagery of vowel /a/. (b) Articulation motor imagery of vowel /o/. (c) Articulation motor imagery of vowel /u/. (d) Articulation motor imagery of vowel /i/.

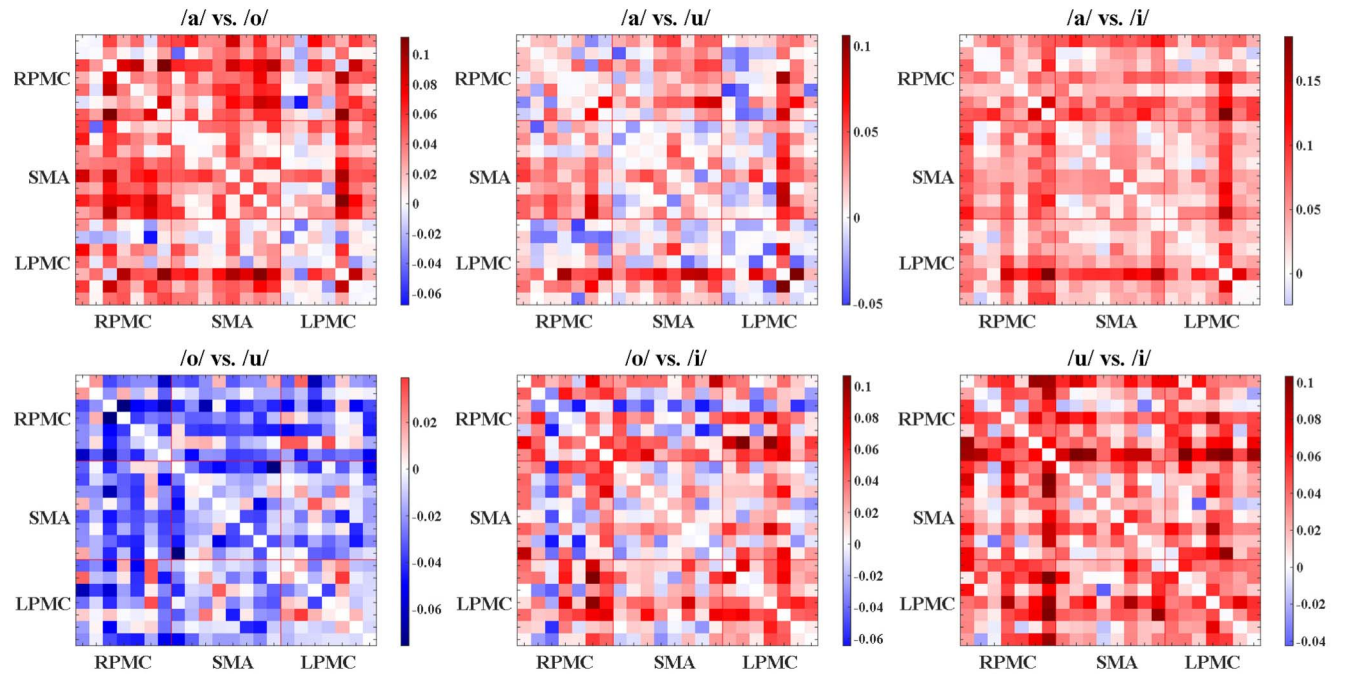

Fig. 7. Paired functional connection strength difference between two vowels in the 0-10 s time window.

vowels. In the 0-2.5 s time window, most connection strengths of vowels /a/, /u/ and /o/ are higher than those of vowel /i/ and many of connection strengths of those vowels are significantly higher than those of vowel /i/.

C. Classification Results

The best classification accuracy was obtained according to the feature selection process. Tables I and II describe the subject-dependent classification accuracies of the connection features extracted in the 0-10 s and 0-2.5 s time windows for each subject, respectively. The classification performances of SVM (Linear, Gaussian, and Polynomial kernels), LDA, and KNN were compared. As shown in Tables I and II, both the average subject-dependent multiclass and binary classification accuracies for LDA in both the 0-10 s and 0-2.5 s time windows are higher than the other machine learning algorithms.

Table I shows the classification accuracies in the 0-10 s time window. For the multiclass (four-class) classification of LDA in the 0-10 s time window, the mean classification accuracy reached 40.9%. The classification accuracy of each subject was higher than the random chance level (i.e., 25%). The average classification accuracies of all pairwise comparisons between two vowels were calculated for each subject as the binary classification accuracy. For the binary classification of LDA in the 0-10 s time window, the average classification accuracy was 71.6% which exceeded the marginal classification accuracy rate (i.e., 70%) [33] that determines whether a binary BCI system could be used for practical application. The classification accuracy of each subject exceeded the random chance level (i.e., 50%).

Table II lists the classification accuracies using early information in the 0-2.5 s time window. For the multiclass classification of LDA in the 0-2.5 s time window, the mean classification accuracy was 39.5%. Each value of the classification accuracy was higher than the random chance level (25%). For binary classification, the average classification accuracy

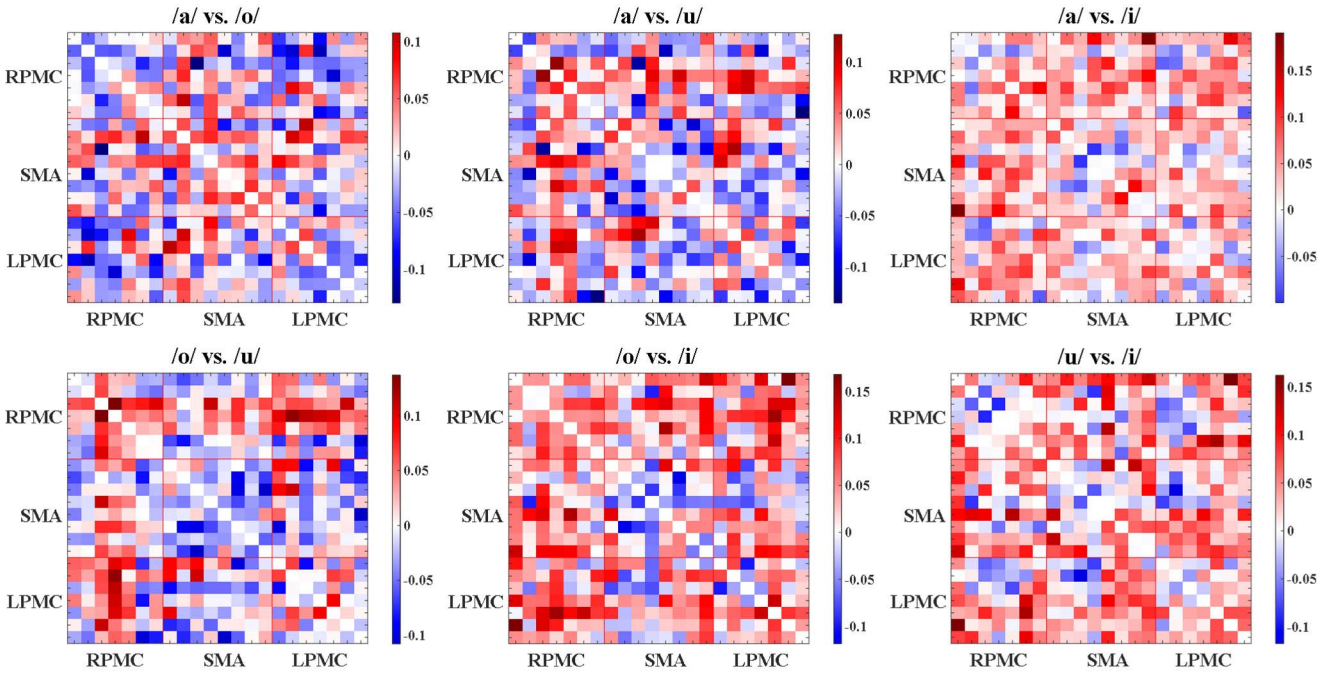

Fig. 8. Paired functional connection strength difference between two vowels in 0-2.5 s time window.

TABLE I
THE CLASSIFICATION ACCURACIES (%) OF CONNECTION FEATURES IN 0-10 S TIME WINDOW

	Four class classification					Binary classification				
	LDA	SVM (Linear)	SVM (Gaussian)	SVM (Polynomial)	KNN	LDA	SVM (Linear)	SVM (Gaussian)	SVM (Polynomial)	KNN
Subject 1	37.50	30.00	28.75	28.75	30.00	63.75	64.17	64.58	63.75	65.42
Subject 2	40.00	40.00	40.00	40.00	42.50	72.08	69.58	66.67	68.75	67.92
Subject 3	33.75	38.75	35.00	38.75	36.25	73.33	76.67	70.00	69.58	67.92
Subject 4	37.50	37.50	38.75	35.00	31.25	72.08	67.92	66.25	68.33	71.25
Subject 5	43.75	38.75	35.00	32.50	41.25	70.83	63.75	58.75	62.50	64.58
Subject 6	42.50	43.75	30.00	30.00	33.75	62.92	59.58	59.58	60.83	64.17
Subject 7	38.75	36.25	35.00	32.50	28.75	72.08	71.25	71.25	69.17	66.67
Subject 8	43.75	41.25	38.75	38.75	32.50	79.58	73.33	73.33	72.50	71.25
Subject 9	41.25	36.25	32.50	37.50	37.50	72.92	69.58	69.58	68.75	71.67
Subject 10	40.00	38.75	41.25	38.75	36.25	75.00	71.25	66.67	69.17	64.17
Subject 11	42.50	45.00	45.00	37.50	41.25	77.08	75.00	71.67	72.50	68.33
Subject 12	42.50	46.25	41.25	40.00	42.50	75.00	74.17	72.92	72.50	73.75
Subject 13	40.00	38.75	40.00	36.25	35.00	75.83	72.08	71.67	70.00	71.25
Subject 14	51.25	47.50	52.50	48.75	47.50	73.75	72.50	74.58	75.83	77.92
Subject 15	46.25	23.75	23.75	27.50	28.75	61.67	60.00	55.00	60.83	57.08
Subject 16	42.50	46.25	53.75	46.25	45.00	77.92	72.08	72.08	72.92	75.42
Subject 17	32.50	28.75	30.00	27.50	27.50	65.00	65.00	60.00	59.58	65.00
Subject 18	40.00	38.75	35.00	36.25	37.50	67.50	58.75	62.50	65.00	64.58
Mean	40.90	38.68	37.57	36.25	36.39	71.57	68.70	67.06	67.92	68.24
STD	4.30	6.31	7.73	5.86	5.99	5.32	5.55	5.83	4.77	4.97

was 71.9%, and the classification accuracy of each subject exceeded 50%.

As can be seen from Tables I and II, although connectivity features were extracted from the reduced (0-2.5 s) time window, both the multiclass and binary classification accuracies based on those features were comparable with that in the 0-10 s time window. Table SII (in Supplementary Material) depicts the significance level p-value of paired t-test among the classification accuracies acquired through different time windows for different machine learning methods. The differences between multiclass/binary classification accuracies acquired through different time windows for all machine learning methods are not significant. Those results might indicate that the functional connectivity features extracted

TABLE II
THE CLASSIFICATION ACCURACIES (%) OF CONNECTION FEATURES IN 0-2.5 S TIME WINDOW

	Four class classification					Binary classification				
	LDA	SVM (Linear)	SVM (Gaussian)	SVM (Polynomial)	KNN	LDA	SVM (Linear)	SVM (Gaussian)	SVM (Polynomial)	KNN
Subject 1	30.00	27.50	28.75	33.75	26.25	69.17	64.58	57.92	62.08	67.92
Subject 2	33.75	32.50	31.25	33.75	36.25	74.58	68.33	71.25	65.83	70.42
Subject 3	43.75	41.25	46.25	47.50	43.75	78.33	78.33	75.83	75.42	71.67
Subject 4	37.50	36.25	36.25	41.25	40.00	70.00	67.92	67.50	65.00	63.75
Subject 5	46.25	41.25	41.25	40.00	45.00	72.08	69.17	69.17	71.25	72.92
Subject 6	41.25	37.50	38.75	40.00	38.75	70.42	63.75	60.42	66.67	65.00
Subject 7	36.25	37.50	30.00	33.75	35.00	63.75	56.67	54.17	59.58	62.08
Subject 8	55.00	52.50	58.75	51.25	50.00	78.75	80.42	80.42	77.92	80.42
Subject 9	33.75	32.50	32.50	33.75	36.25	65.83	63.33	61.25	63.75	64.17
Subject 10	40.00	37.50	35.00	32.50	33.75	73.75	68.75	66.67	65.83	64.58
Subject 11	36.25	28.75	27.50	32.50	35.00	69.58	70.00	69.17	69.58	69.58
Subject 12	42.50	41.25	31.25	38.75	28.75	73.33	65.42	65.00	65.00	66.25
Subject 13	36.25	37.50	36.25	35.00	35.00	68.33	67.50	61.67	64.17	64.17
Subject 14	47.50	45.00	46.25	46.25	43.75	72.08	72.92	72.08	72.92	71.67
Subject 15	37.50	38.75	42.50	36.25	33.75	75.83	71.25	72.08	70.42	67.08
Subject 16	41.25	43.75	35.00	45.00	36.25	77.08	76.25	73.75	75.00	72.92
Subject 17	40.00	32.50	26.25	35.00	33.75	68.33	68.75	62.50	64.58	66.25
Subject 18	32.50	32.50	35.00	40.00	31.25	72.08	68.75	64.17	65.83	64.17
Mean	39.51	37.57	36.60	38.68	36.81	71.85	69.00	66.94	67.82	68.06
STD	6.08	6.15	8.10	5.70	5.93	4.14	5.63	6.72	5.02	4.62

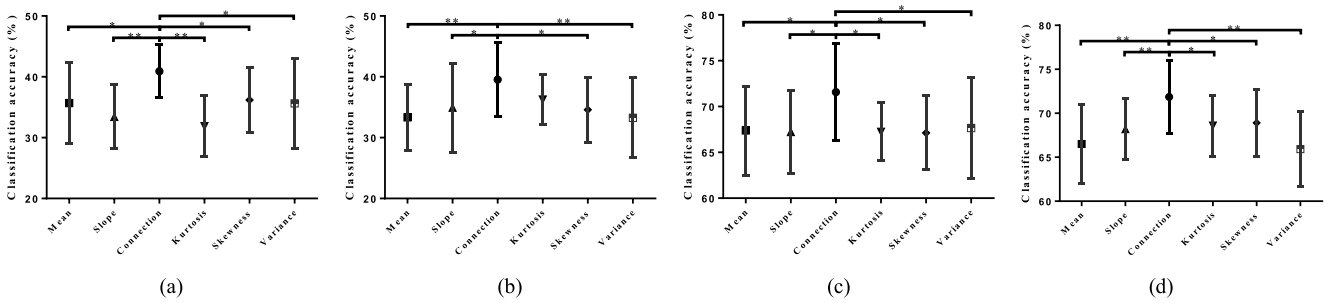

Fig. 9. The multiclass and binary classification accuracies (%) of LDA in 0-10 s and 0-2.5 s time windows. (a) The four-class classification accuracies in 0-10 s time window. (b) The four-class classification accuracies in 0-2.5 s time window. (c) The binary classification accuracies in 0-10 s time window. (d) The binary classification accuracies in 0-2.5 s time window. * and ** mean the *p*-value of paired t-test below 0.05 and 0.005, respectively.

from the 0-2.5 s time window can provide as much differentiated information as that extracted from the 0-10 s time window.

Figure 9 shows the multiclass and binary classification accuracies of connection and activation features using LDA in 0-10 s and 0-2.5 s time windows. No matter in multiclass setting or binary setting, the average classification accuracy of each type of activation features is lower than that of connection features. Paired t-test was conducted among the classification accuracies acquired through different features in different time windows. The decoding performance of functional connectivity features is significantly higher than that of most temporal features. Those results may suggest that connection features could provide more distinguishable information than activation features.

D. The Role of Different Connections in Providing Discriminative Information

Figure 10 shows the F-score matrices in both the 0-10 s and 0-2.5 s time windows. F-score was the measure of each functional connection feature in providing the discrimination power for classifying different articulation motor imagery tasks. According to the distributions of the F-score matrices in both the 0-10 s and 0-2.5 s time windows, the connections with high discrimination power mainly concentrated on the synchronization between the left PMC and the right PMC or SMA and between channels located on the left PMC. Figure 11 shows the 20 most selected connection features of HbO signal in different time windows. These selected features indicate that the features with strong discriminative power usually were selected to classify and the connections

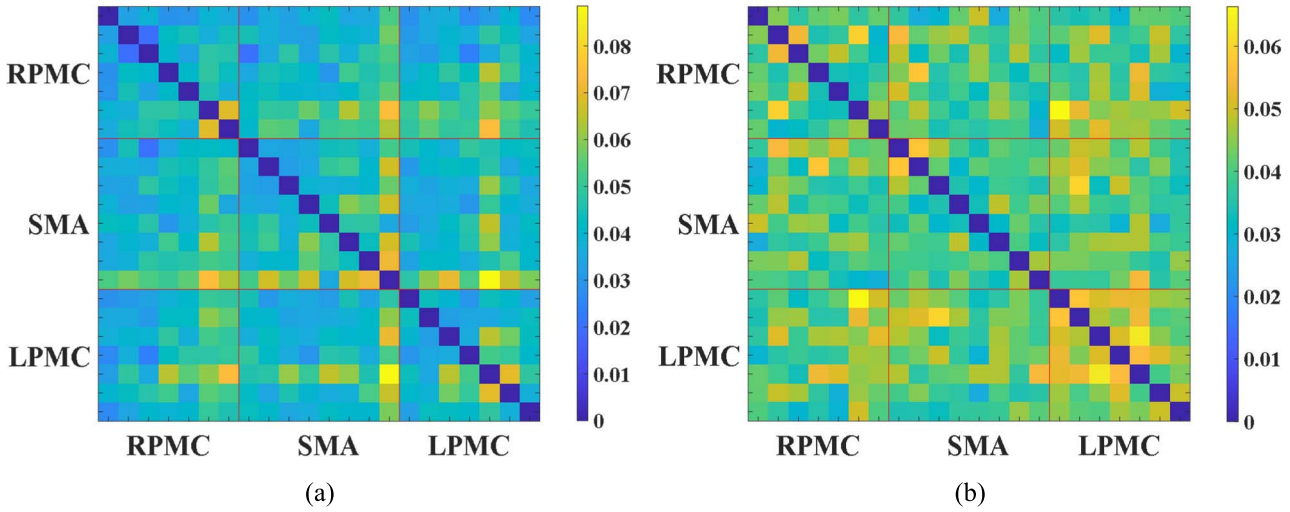

Fig. 10. *F*-score matrixes of connection features in (a) 0-10 s time window, (b) 0-2.5 s time window.

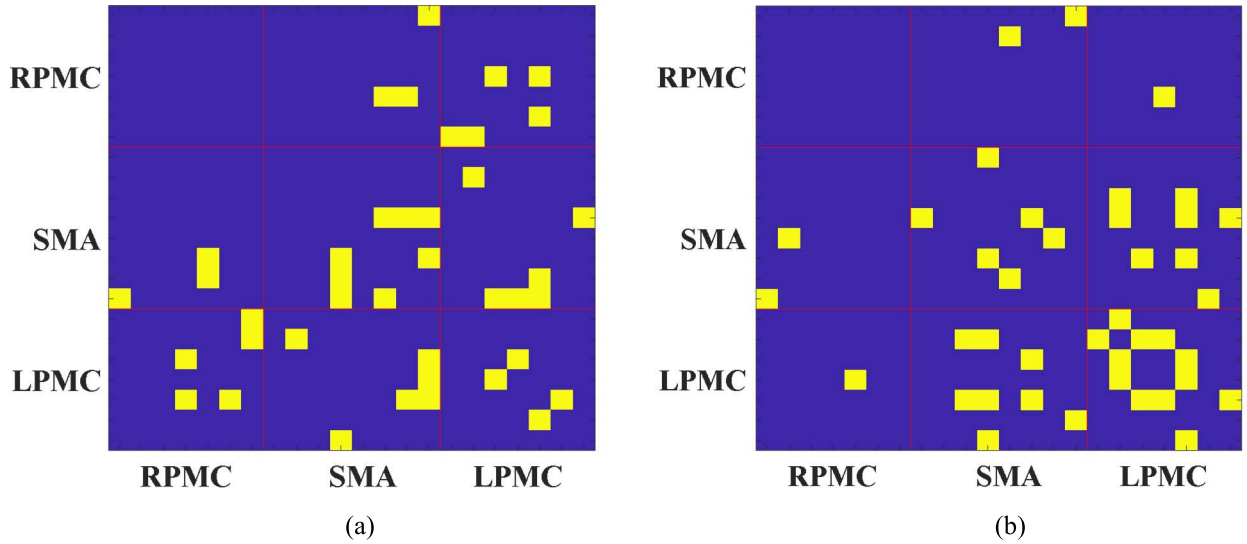

Fig. 11. Most selected features of HbO signals in different time windows. (a) 0-10 s time window, (b) 0-2.5 s time window.

which concentrated on the left PMC were usually selected as features. Those results suggested that the left PMC play a more important role in identifying different articulation motor imagery tasks than the right PMC.

IV. Discussion

This study aimed to establish an fNIRS-based BCI to decode articulation motor imagery of four vowels. Relevant features extracted from the brain functional network in the motor cortex can efficiently decode different vowels. Features extracted from the neural network in the reduced (0-2.5 s) time window can provide much information to distinguish different vowel imagery tasks no less than that in the long (0-10 s) time window. Those results suggest that speech imagery BCI based on fNIRS can be further optimized in terms of feature extraction and command generation time reduction.

Figure 3 shows the average value of HbO concentration changes of 22 channels for different tasks. As shown in Fig. 3, two large bilateral clusters of high activations located on the PMC and a small parietal cluster of high activations located on the SMA were found. Those brain areas with high activations revealed by the activation analysis were indicated heavily involved in articulation motor imagery. These results are consistent with the previous fMRI study [34] suggesting that the bilateral and parietal motor areas are involved in controlling lip and jaw movements. As shown in Figs. 5 and 6, the connection coefficients of interaction between the left PMC and other brain regions and between channels located on the left PMC are higher than that between the right PMC with other neural areas and between channels located on the right PMC. Figure 10 shows that the connections with strong discriminative information mainly concentrate to the left PMC in both the 0-10 s and 0-2.5 s time windows. Given that

the left PMC plays a more important role in motor imagery of articulation than the right PMC, this finding is in line with the previous finding [34] that a significant left-favoring hemispheric activation asymmetry was found in lip and jaw movement control.

As can be seen from Fig. 3, the activation level of HbO of imagining the articulation movements of vowel /o/ is higher than that of other vowels. Fig. 4 shows the activation levels of vowel /o/ are significantly higher in some brain areas than that of the other vowels. The articulation movements of vowel /o/ contained jaw opening and lip rounding which is more complex than that of other vowels. These results suggest that different complexities of the articulation motor imagery tasks lead to different activation levels of HbO. This phenomenon has also been shown in previous studies. HbO changes in response to finger-tapping tasks with different complexities can be detected by fNIRS [35]. Holper et al. [36] also found fNIRS can be used to differentiate motor imagery tasks with different complexities (a simple finger-tapping imagery task and a complex sequential finger-tapping imagery task), and the average classification accuracy across 12 subjects reached 81%. The impact of task complexity on brain activities also has been examined by other functional neuroimaging modalities. Catalan et al. [37] found higher brain activations for complex finger movements than for simple finger movements using PET. Similar conclusions also have been obtained by another study using PET with different complex hand movement tasks [38]. Furthermore, Kuhtz-Buschbeck et al. [39] found the motor evoked potentials showed a significant rise in the complex imagined finger movements task compared to that in the simple imagined finger movements task using fMRI combined with transcranial magnetic stimulation.

Figure 7 shows that most connection strengths of vowel /a/ are higher than those of the other vowels and most connection strengths of vowel /i/ are lower than those of the other vowels in the 0-10 s time window. As depicted in Fig. 8, most connection strengths of vowel /i/ are lower than those of the other vowels in the 0-2.5 s time window. The strong connection strength of vowel /a/ and weak connection strength of vowel /i/ may be caused by different force levels of jaw opening. Previous studies also have found similar phenomena. Nambu et al. [40] found that different finger-pinch force levels (25, 50, or 75% of the maximum voluntary contraction (MVC)) could be recognized from the activity in the human brain using fNIRS. Zheng et al. [41] explored the functional connectivity between multiple cortical areas during grip force tracking tasks at 25%, 50%, and 75% of MVC with fNIRS and found the functional connectivity between the left prefrontal cortex and the left sensorimotor cortex and between the right prefrontal cortex and the right sensorimotor cortex strengthened with a higher grip force level. Ortega et al. [42] have revealed traces of the brain activity being modulated by the level of hand-specific forces (10, 17.5, or 25% of MVC) using EEG and fNIRS and reconstructed bimanual force trajectories. Andrushko et al. [43] found increases in right-handgrip force resulted in greater ipsilateral sensorimotor activation and greater functional connectivity between hemispheres within the sensorimotor network with 25%, 50%, and 75% of MVC. Furthermore, different imagined force levels also were observed leading to different brain activities and functional connectivity. Mizuguchi et al. [44] investigated the relationship between brain activity and imagined right hand grasping force level (10%, 30%, and 60% of MVC) with fMRI and observed right fronto-parietal activity increased with intensifying of the imagined force level. Yin et al. [45] decoded imagined both force (20/50/80% MVC) and speed (0.5/1/2 Hz) of hand clenching with fNIRS and EEG and the classification accuracy of both force and speed of hand clenching achieved 89% $\pm$ 2%. Yin et al. [46] investigate the brain hemodynamic responses of hand clench force (20/50/80% MVC) and speed (0.5/1/2 Hz) imagery with fNIRS and obtained an average classification accuracy of 78%. Fu et al. [47] explored the relationship between brain activity and imagined hand clenching force (20/50/80% MVC) and speed (0.5/1/2 Hz) with EEG and fNIRS and reported an average 6-class classification accuracy of 74%$\pm$2%. In this study, the force levels of jaw opening are different for the vowel /a/, /o/, and /i/. The force level of jaw opening of /a/ is the largest and the force level of jaw opening of /i/ is the least. Therefore, the connection strengths of different vowels were observed to be varied and the functional connectivity provided a lot of information to classify different vowels. Most connection strengths of vowel /o/ were lower than that of vowel /a/ may be due to the functional connectivity being more sensitive to force level rather than task complexity.

As depicted in Tables I and II, the classification accuracies of features extracted in the 0-2.5 s time window were similar to that in the 0-10 s time window for all the machine learning algorithms used in this study, which indicated that the functional connectivity in the reduced time window could provide as much distinguishable information as that in the long time window. This finding may be caused by the initial dip [48] in the early HbO and HbR responses phase (0-2.5 s time window). Initial dip is defined as an initial small HbR/HbO concentration change related to specific neuronal activity after functional stimulation [49]. The initial dip originated from the early extraction of HbO by locally metabolized neurons from the capillary network. Therefore, it is believed to be a faster and better spatial localizer on the neuronal activity than the hemodynamic response [50]. Because the initial dip might be more sensitive to the neuronal activity than the hemodynamic response, the synchronization of hemodynamic responses in the 0-2.5 s time window (including initial dips) of different brain regions provided as much distinguishable information as that in 0-10 s time window. As shown in Fig. 6, more connections with high F-scores concentrate on the left motor cortex in the 0-2.5 s time window than that in the 0-10 s time window which supports that the connection information extracted from in the 0-2.5 s time window might be more closely related to neuronal activity.

Compared with the existing speech imagery BCIs based on fNIRS [6], [7], [8], [9], this is the first study to classify vowels by decoding the imagery of simplified articulation movements. The existing speech imagery BCIs based on fNIRS mainly focus on using the information of acoustic and semantic differences in the auditory cortex to decode the imagined

speech. The binary classification accuracies of those BCIs approximated or exceeded 70% which is comparable with that of our study. Those results indicated that the imagery of articulation movements could provide as much distinguishable information as acoustic and semantic differences in speech imagery. In the future, the decoding performance of the speech imagery BCIs based on fNIRS might be improved by combining both the articulation motor imagery information with acoustic and semantic information in speech imagery.

V. Conclusion

BCI based on speech imagery can help individuals with very limited or no voluntary movements express their intention in a user-friendly way. An fNIRS-based BCI was presented to decode simplified articulation motor imagery of four vowels. In this work, functional connectivity features extracted from fNIRS signals were used to classify articulation motor imagery of different vowels. In multiclass (four classes) settings, the mean classification accuracies approximated or exceeded 40% in the 0-2.5 s and 0-10 s time windows, respectively. In binary settings, the mean classification accuracies exceeded 70% in both the 0-2.5 s and 0-10 s time windows. These encouraging results demonstrate that connection features extracted from fNIRS signals can effectively differentiate articulation motor imagery of different vowels, and are still valid even in a shorter time window (0-2.5 s), which are significant for reducing command generation time and feature extraction for fNIRS-based BCIs. Further, simplified articulation motor imagery of vowels can be distinguished using the information extracted from the motor cortex, and therefore, speech imagery BCIs based on fNIRS might achieve better performance by combining the acoustic and semantic information with the simplified articulation movements imagery information.

Acknowledgment

The authors thank Prof. Lisheng Xu for writing assistance.

References

[1] L. F. Nicolas-Alonso and J. Gomez-Gil, "Brain computer interfaces, a review," *Sensors*, vol. 12, no. 2, pp. 1211–1279, 2012.

[2] T. Schultz, M. Wand, T. Hueber, D. J. Krusienski, C. Herff, and J. S. Brumberg, "Biosignal-based spoken communication: A survey," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 25, no. 12, pp. 2257–2271, Dec. 2017.

[3] K.-S. Hong, U. Ghafoor, and M. J. Khan, "Brain–machine interfaces using functional near-infrared spectroscopy: A review," *Artif. Life Robot.*, vol. 25, no. 2, pp. 204–218, May 2020.

[4] R. Abiri, S. Borhani, Y. Jiang, and X. Zhao, "Decoding attentional state to faces and scenes using EEG brainwaves," *Complexity*, vol. 2019, pp. 1–10, Feb. 2019.

[5] F. Herold, P. Wiegel, F. Scholkmann, and N. Müller, "Applications of functional near-infrared spectroscopy (fNIRS) neuroimaging in exercise-cognition science: A systematic, methodology-focused review," *J. Clin. Med.*, vol. 7, no. 12, p. 466, 2018.

[6] K. Tanino, H. Miura, N. Matsuda, and H. Taki, "The analysis of the brain state measuring by NIRS-based BMI in answering yes-no questions," *Proc. Comput. Sci.*, vol. 60, pp. 1233–1239, Jan. 2015.

[7] H.-J. Hwang et al., "Toward more intuitive brain–computer interfacing: Classification of binary covert intentions using functional near-infrared spectroscopy," *J. Biomed. Opt.*, vol. 21, no. 9, Apr. 2016, Art. no. 091303.

[8] A. R. Sereshkeh, R. Yousefi, A. T. Wong, and T. Chau, "Online classification of imagined speech using functional near-infrared spectroscopy signals," *J. Neural Eng.*, vol. 16, no. 1, Feb. 2019, Art. no. 016005.

[9] G. Gallegos-Ayala, A. Furdea, K. Takano, C. A. Ruf, H. Flor, and N. Birbaumer, "Brain communication in a completely locked-in patient using bedside near-infrared spectroscopy," *Neurology*, vol. 82, no. 21, pp. 1930–1932, 2014.

[10] M. Roth et al., "Possible involvement of primary motor cortex in mentally simulated movement: A functional magnetic resonance imaging study," *NeuroReport*, vol. 7, no. 7, pp. 1280–1284, May 1996.

[11] X. Wu, W.-L. Zheng, Z. Li, and B.-L. Lu, "Investigating EEG-based functional connectivity patterns for multimodal emotion recognition," *J. Neural Eng.*, vol. 19, no. 1, Feb. 2022, Art. no. 016012.

[12] C.-R. Phang and L.-W. Ko, "Global cortical network distinguishes motor imagination of the left and right foot," *IEEE Access*, vol. 8, pp. 103734–103745, 2020.

[13] P. Li et al., "EEG based emotion recognition by combining functional connectivity network and local activations," *IEEE Trans. Biomed. Eng.*, vol. 66, no. 10, pp. 2869–2881, Oct. 2019.

[14] L. Gu, Z. Yu, T. Ma, H. Wang, Z. Li, and H. Fan, "EEG-based classification of lower limb motor imagery with brain network analysis," *Neuroscience*, vol. 436, pp. 93–109, Jun. 2020.

[15] Q. Ai et al., "Feature extraction of four-class motor imagery EEG signals based on functional brain network," *J. Neural Eng.*, vol. 16, no. 2, Apr. 2019, Art. no. 026032.

[16] Z. S. Saad, K. M. Ropella, R. W. Cox, and E. A. DeYoe, "Analysis and use of FMRI response delays," *Hum. Brain Mapping*, vol. 13, no. 2, pp. 74–93, 2001.

[17] K.-S. Hong, U. Ghafoor, and M. J. Khan, "Brain–machine interfaces using functional near-infrared spectroscopy: A review," *Artif. Life Robot.*, vol. 25, no. 2, pp. 204–218, May 2020.

[18] A. Zafar and K.-S. Hong, "Detection and classification of three-class initial dips from prefrontal cortex," *Biomed. Opt. Exp.*, vol. 8, no. 1, pp. 367–383, 2017.

[19] R. Li, T. Potter, W. Huang, and Y. Zhang, "Enhancing performance of a hybrid EEG-fNIRS system using channel selection and early temporal features," *Frontiers Hum. Neurosci.*, vol. 11, p. 462, Sep. 2017.

[20] M. J. Khan and K.-S. Hong, "Hybrid EEG–fNIRS-Based eight-command decoding for BCI: Application to quadcopter control," *Frontiers Neurorobotics*, vol. 11, p. 6, Feb. 2017.

[21] A. Zafar and K.-S. Hong, "Neuronal activation detection using vector phase analysis with dual threshold circles: A functional near-infrared spectroscopy study," *Int. J. Neural Syst.*, vol. 28, no. 10, Dec. 2018, Art. no. 1850031.

[22] D. Tsuzuki, V. Jurcak, A. K. Singh, M. Okamoto, E. Watanabe, and I. Dan, "Virtual spatial registration of stand-alone fNIRS data to MNI space," *NeuroImage*, vol. 34, no. 4, pp. 1506–1518, Feb. 2007.

[23] H. Gray, *Anatomy of the Human Body*. Philadelphia, PA, USA: Lea & Febiger, 1878.

[24] J. C. Ye, S. Tak, K. E. Jang, J. Jung, and J. Jang, "NIRS-SPM: Statistical parametric mapping for near-infrared spectroscopy," *NeuroImage*, vol. 44, no. 2, pp. 428–447, 2009.

[25] S. Jahani, S. K. Setarehdan, D. A. Boas, and M. A. Yücel, "Motion artifact detection and correction in functional near-infrared spectroscopy: A new hybrid method based on spline interpolation method and Savitzky–Golay filtering," *Neurophotonics*, vol. 5, no. 1, Feb. 2018, Art. no. 015003.

[26] X. Si, S. Li, S. Xiang, J. Yu, and D. Ming, "Imagined speech increases the hemodynamic response and functional connectivity of the dorsal motor cortex," *J. Neural Eng.*, vol. 18, no. 5, Oct. 2021, Art. no. 056048.

[27] M. A. Yaqub, S.-W. Woo, and K.-S. Hong, "Effects of HD-tDCS on resting-state functional connectivity in the prefrontal cortex: An fNIRS study," *Complexity*, vol. 2018, pp. 1–13, Nov. 2018.

[28] K. M. Arun, K. A. Smitha, P. N. Sylaja, and C. Kesavadas, "Identifying resting-state functional connectivity changes in the motor cortex using fNIRS during recovery from stroke," *Brain Topography*, vol. 33, no. 6, pp. 710–719, Nov. 2020.

[29] A. Eken, "Assessment of flourishing levels of individuals by using resting-state fNIRS with different functional connectivity measures," *Biomed. Signal Process. Control*, vol. 68, Jul. 2021, Art. no. 102645.

[30] H.-J. Hwang, J.-H. Lim, D.-W. Kim, and C.-H. Im, "Evaluation of various mental task combinations for near-infrared spectroscopy-based brain-computer interfaces," *J. Biomed. Opt.*, vol. 19, no. 7, Jul. 2014, Art. no. 077005.

[31] Z. Guo and F. Chen, "Idle-state detection in motor imagery of articulation using early information: A functional near-infrared spectroscopy study," *Biomed. Signal Process. Control*, vol. 72, Feb. 2022, Art. no. 103369.
[32] D. Malonek and A. Grinvald, "Interactions between electrical activity and cortical microcirculation revealed by imaging spectroscopy: Implications for functional brain mapping," *Science*, vol. 272, no. 5261, pp. 551–554, Apr. 1996.
[33] A. Kubler, V. K. Mushahwar, L. R. Hochberg, and J. P. Donoghue, "BCI meeting 2005-workshop on clinical issues and applications," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 14, no. 2, pp. 131–134, Jun. 2006.
[34] K. Grabski et al., "Functional MRI assessment of orofacial articulators: Neural correlates of lip, jaw, larynx, and tongue movements," *Hum. Brain Mapping*, vol. 33, no. 10, pp. 2306–2321, Oct. 2012.
[35] L. Holper, M. Biallas, and M. Wolf, "Task complexity relates to activation of cortical motor areas during uni- and bimanual performance: A functional NIRS study," *NeuroImage*, vol. 46, no. 4, pp. 1105–1113, Jul. 2009.
[36] L. Holper and M. Wolf, "Single-trial classification of motor imagery differing in task complexity: A functional near-infrared spectroscopy study," *J. Neuroeng. Rehabil.*, vol. 8, no. 1, pp. 1–13, 2011.
[37] M. J. Catalan, M. Honda, R. A. Weeks, L. G. Cohen, and M. Hallett, "The functional neuroanatomy of simple and complex sequential finger movements: A PET study," *Brain*, vol. 121, pp. 253–264, 1998.
[38] R. Kawashima et al., "Regional cerebral blood flow changes in human brain related to ipsilateral and contralateral complex hand movements—A PET study," *Eur. J. Neurosci.*, vol. 10, no. 7, pp. 2254–2260, Jul. 1998.
[39] J. P. Kuhtz-Buschbeck, C. Mahnkopf, C. Holzknecht, H. Siebner, S. Ulmer, and O. Jansen, "Effector-independent representations of simple and complex imagined finger movements: A combined fMRI and TMS study," *Eur. J. Neurosci.*, vol. 18, no. 12, pp. 3375–3387, Dec. 2003.
[40] I. Nambu, R. Osu, M.-A. Sato, S. Ando, M. Kawato, and E. Naito, "Single-trial reconstruction of finger-pinch forces from human motor-cortical activation measured by near-infrared spectroscopy (NIRS)," *NeuroImage*, vol. 47, no. 2, pp. 628–637, Aug. 2009.
[41] X. Zheng et al., "Detection of functional connectivity in the brain during visuo-guided grip force tracking tasks: A functional near-infrared spectroscopy study," *J. Neurosci. Res.*, vol. 99, no. 4, pp. 1108–1119. 2021.
[42] P. Ortega and A. A. Faisal, "Deep learning multimodal fNIRS and EEG signals for bimanual grip force decoding," *J. Neural Eng.*, vol. 18, no. 4, Aug. 2021, Art. no. 0460e6.
[43] J. W. Andrushko et al., "High force unimanual handgrip contractions increase ipsilateral sensorimotor activation and functional connectivity," *Neuroscience*, vol. 452, pp. 111–125, Jan. 2021.
[44] N. Mizuguchi, H. Nakata, and K. Kanosue, "Activity of right premotor-parietal regions dependent upon imagined force level: An fMRI study," *Frontiers Human Neurosci.*, vol. 8, p. 810, Oct. 2014.
[45] X. Yin 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," *J. Neural Eng.*, vol. 12, no. 3, Jun. 2015, Art. no. 036004.
[46] X. Yin et al., "NIRS-based classification of clench force and speed motor imagery with the use of empirical mode decomposition for BCI," *Med. Eng. Phys.*, vol. 37, no. 3, pp. 280–286, Mar. 2015.
[47] Y. Fu, X. Xiong, C. Jiang, B. Xu, Y. Li, and H. Li, "Imagined hand clenching force and speed modulate brain activity and are classified by NIRS combined with EEG," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 9, pp. 1641–1652, Sep. 2017.
[48] K.-S. Hong and A. Zafar, "Existence of initial dip for BCI: An illusion or reality," *Frontiers Neurorobot.*, vol. 12, p. 69, Oct. 2018.
[49] R. D. Frostig, E. E. Lieke, D. Y. Ts'o, and A. Grinvald, "Cortical functional architecture and local coupling between neuronal activity and the microcirculation revealed by in vivo high-resolution optical imaging of intrinsic signals," *Proc. Nat. Acad. Sci. USA*, vol. 87, no. 16, pp. 6082–6086, Aug. 1990.
[50] I. Vanzetta and A. Grinvald, "Coupling between neuronal activity and microcirculation: Implications for functional brain imaging," *HFSP J.*, vol. 2, no. 2, pp. 79–98, Apr. 2008.