A Novel Neurorehabilitation Prognosis Prediction Modeling on Separated Left-Right Hemiplegia Based on Brain-Computer Interfaces Assisted Rehabilitation

Zhimin Shao, Weibei Dou, Di Ma, Xiaoxue Zhai, Quan Xu, and Yu Pan

Abstract—**It is essential for neuroscience and clinic to estimate the influence of neuro-intervention after brain damage. Most related studies have used Mirrored Contralesional-Ipsilesional hemispheres (MCI) methods flipping the axial neuroimaging on the x-axis in prognosis prediction. But left-right hemispheric asymmetry in the brain has become a consensus. MCI confounds the intrinsic brain asymmetry with the asymmetry caused by unilateral damage, leading to questions about the reliability of the results and difficulties in physiological explanations. We proposed the Separated Left-Right hemiplegia (SLR) method to model left and right hemiplegia separately. Two pipelines have been designed in contradistinction to demonstrate the validity of the SLR method, including MCI and removing intrinsic asymmetry (RIA) pipelines. A patient dataset with 18 left-hemiplegic and 22 right-hemiplegic stroke patients and a healthy dataset with 40 subjects, age- and sex-matched with the patients, were selected in the experiment. Blood-Oxygen Level-Dependent MRI and Diffusion Tensor Imaging were used to build brain networks whose nodes were defined by the Automated Anatomical Labeling atlas. We applied the same statistical and machine learning framework for all pipelines, logistic regression, artificial neural network, and support vector machine for classifying the patients who are significant or non-significant responders to brain-computer interfaces assisted training and optimal subset regression, support vector regression for predicting post-intervention outcomes. The SLR pipeline showed 5-15% improvement in accuracy and at least 0.1 upgrades in R^2, revealing common and unique recovery mechanisms after left and right strokes and helping clinicians make rehabilitation plans.**

Index Terms—**Brain-computer interface, BCI, machine learning, brain asymmetry, MRI.**

I. Introduction

SINCE the 20th century, neurorehabilitation scientists have realized that injured brain tissue has the potential for regeneration and plasticity [1]. Neurorehabilitation intervention studies try to help patients with brain disorders recover from dysfunction. For instance, constraint-induced movement therapy, Transcranial magnetic stimulation, and Brain-computer interfaces (BCI) are used to rehabilitate motor recovery [2], [3], [4], [5] after stroke. Typically, a rehabilitation program lasts for weeks or months. Due to limited medical resources, the prognostic prediction of treatment performance is necessary for the clinic to make rehabilitation plans.

Functional connectivity based on resting-state functional magnetic resonance imaging (rs-fMRI) measures the blood-oxygen-level-dependent (BOLD) signal's temporal correlation across brain regions without imposed tasks [6]. Diffusion tensor imaging (DTI), one kind of Diffusion MRI, provides insights into the tissue structure [7]. MRI modalities can be applied to disease diagnosis and prognosis based on quantitative imaging biomarkers [8], [9], [10], [11]. However, biomarkers were mainly derived from a single brain region, and the connections were ignored across regions and modalities, remaining unsatisfactory in predicting post-interventional outcomes at the individual level [12], [13], [14], [15]. The graph theory-based approach analyzes the whole brain as a complex network, where brain regions are regarded as nodes, and edges indicate the relationships between nodes [16], [17]. The network demonstrates the relationship between brain structure and function and the characteristics of the whole brain in high-dimensional feature space.

Machine learning methods are up-and-coming for the study of high-dimensional neuroimaging data. Reference [18] summarized the typical classification/prediction pipeline for post-stroke recovery. Due to the small sample size, researchers commonly mirrored the imaging along the mid-sagittal axis so that all tracts were in the same hemisphere [13], [15]. Therefore, the issue of left-right brain hemispheres was converted into that of contra-ipsilesional hemispheres. Nevertheless, hemispheric asymmetry has become a consensus for brain structure and function [19], [20]. Regarding stroke patients, researchers aim to distinguish their difference from healthy controls and then predict the corresponding recovery potential according to clinical and imaging metrics. The intrinsic left-right asymmetry confuses the difference caused by impairment and recovery.

We propose modeling the left and right hemiplegia groups separately, which is called SLR modeling Pipeline in this paper to sidestep the effect of left-right hemispheric asymmetry. Two alternative modeling pipelines – the mirrored contra-ipsilesional hemispheres (MCI) Pipeline and the removing intrinsic asymmetry (RIA) Pipeline are designed as a comparison to verify the effectiveness of SLR. The subjects and BCI rehabilitation are presented in Section II. In Section III, some evidence shows that the proposed SLR pipeline may improve the accuracy of the prognostic classification and prediction of the treatment performance, which is explained in detail in Sectionn IV.

II. MATERIALS AND METHODS

A. Participants

The study protocol was approved by the Ethics Committee of Beijing Tsinghua Changgung Hospital, conducted in accordance with the tenets of the Declaration of Helsinki (No. 18172-0-02), and registered in the Chinese Clinical Trial Registry (No. ChiCTR1900022128). All patients provided written informed consent prior to study participation.

A patient dataset with 18 left-hemiplegic and 22 right-hemiplegic stroke patients and a healthy dataset with 40 subjects, age- and sex-matched with the patients, were used in the study. Every five years was an age group, and subjects in the same group were considered age-matched. All stroke patients were selected from hospitalized in Beijing Tsinghua Changgung Hospital from March 2018 to October 2019. The diagnosis was in line with the standard consensus of Clinical research on acute stroke in China and confirmed by brain MRI or CT examination. Five criteria are considered for study inclusion: (i) age 18-75; (ii) no significant cognitive impairment (score of simple mental state examination $\geq$ 21); (iii) first-time stroke and time since onset $>$ 1 month and $\leq$ 6 months; (iv) moderate to severe upper limb motor dysfunction (Brunnstrom stage $\leq$ Stage IV [21]); (v) modified Ashworth $<$ III [22]. Detailed demographic and clinical information are listed in Table I.

The 40 healthy control subjects were selected from the HCP dataset (including HCP 1.0, 1200 participants, and HCP 2.0, 1300 participants) [23], with age- and gender-matched to the 40 patients, for performing the symmetrical feature selection procedure.

TABLE I
PATIENTS' INFORMATION

	Significant responders (Mean ± SD)	Non-significant responders (Mean ± SD)	P
Gender (male/female)	16 / 7	11 / 6	0.7533
Hemiplegia (left/right)	10 / 13	8 / 9	0.8750
Age	57.09 ± 10.22	56.47 ± 12.25	0.8633
Course of disease	2.00 ± 1.04	2.59 ± 1.23	0.1104
Pre FMA-total	22.61 ± 13.38	14.53 ± 18.26	0.1142
Pre FMA-SE	16.57 ± 7.20	10.18 ± 10.11	0.0248
Pre FMA-WH	4.78 ± 5.52	3.82 ± 7.31	0.6390
Post FMA-total	38.09 ± 12.53	19.76 ± 17.83	0.0005*
Post FMA-SE	23.35 ± 6.15	14.71 ± 9.57	0.0013*
Post FMA-WH	12.17 ± 6.07	4.06 ± 7.32	0.0005*
Lesion Type (isch/hemo)	18 / 5	10 / 7	0.1941

SD: standard deviation; Pre: pre-interventional; Post: post-interventional; isch: ischemic stroke; hemo: hemorrhagic stroke.
* Statistical result of two groups in different treatments at significance level 0.01.

Please see Supplementary Table S1 for detailed information about each subject and matched healthy controls. Supplementary Table S2 and Supplementary Fig.S1 show the lesion location and size among the individuals.

B. Treatment

Each patient received a 20-day BCI-assisted rehabilitation training program [24], which is an exoskeleton hand controlled by BCI signal to assist the paretic hand exercises. The BCI recognizes the grasping or releasing motor imagery signal of the patient's Electroencephalography (EEG) when the imagination according to the guidance shown on screen and voice prompts (10 sets of 10 imagery tasks in each session, together with passive flexion and extension of fingers through the exoskeleton, 1 min rest between sets, total lasting 1 hour). There are five rehabilitation training sessions per week for four weeks. During the training, patients were instructed to avoid blinking, coughing, chewing, and head and body movements.

C. Measurement

Before and after the training, Fugl–Meyer's Assessment of the upper extremity [25] was used to evaluate the motor function of the upper limb, including shoulder, elbow, wrist, hand, and coordination/speed. By summing the five items, we got the total score, denoted as FMA-total (66 points) in this paper. The sum of shoulder and elbow scores was the FMA upper arm score, defined as FMA-SE (30 points). The sum of wrist and hand scores was the FMA wrist score, defined as FMA-WH (24 points). The higher the score, the stronger the function. Two trained rehabilitation therapists completed the assessment who were limited to know the treatment plans.

Brain MRI scanning was performed with a GE 3.0T MR scanner (DISCOVERY MR750 model; General Electric American, Waukesha, WI, USA). T1-weighted images (T1), resting-state BOLD-fMRI, and DTI were acquired using a standard 32-channel head coil. fMRI data of BOLD are scanned using "Ax-BOLD-rest" series with a gradient echo-planar imaging sequence (repetition time [TR] = 2000 ms, echo time [TE] = 30 ms, flip angle [FA] = 90°, pixel space = 3.5 mm^2, slice

thickness = 3.5 mm, spacing between slices = 4 mm, acquisition matrix = [64, 0, 0, 64], equivalent to in-plane resolution = 64 × 64, reconstruction diameter = 224 mm, 34 axial slices, and 240 temporal positions). T1-weighted images are scanned using the "Sag 3D T1BRAVO" series (TR = 8.21 ms, TE = 3.18 ms, FA = 8°, voxel space = 1 mm^3, spacing between slices = 1 mm, acquisition matrix = [0, 256, 256, 0], equivalent to 256 axial slices and 256 coronal slices). The sagittal slice number depended on the head size of each subject, ranging from 156 to 174 mm, corresponding to the 40 subjects in this study. The reconstruction diameter was 256 mm. Parameters for DTI images were "Ax_DTI_24_Directions" series with TR = 14.5 s, TE = 112 ms, FA = 90°, slice thickness = 2 mm, spacing between slices = 2 mm.

D. Labeling of the Patient Dataset

Patients were divided into a significant responder group (n = 23) and a non-significant responder group (n = 17) according to the definition of the minimum clinically important difference, based on an improvement of $\geq$ 2 points in the FMA-WH score after training [26].

E. Pipelines

As shown in Fig. 1, the proposed SLR modeling Pipeline consists of 5 steps: preprocessing of whole-brain neuroimaging data, brain atlas registration, whole-brain feature extraction, and left / right classification modeling or prediction modeling. Atlases are used to define brain regions. The features of the whole brain include intra- and inter-region and could be obtained by building brain networks with brain regions as nodes or other methods. In SLR Pipeline, we separate patients into two groups beforehand–left and right, referring to the hemiplegia side, and directly carry out the procedure of classifying and predicting based on extracted whole-brain features to achieve the two goals: to classify patients into two classes–significant responders versus non-significant responders to the treatment and to predict stroke patients' function after rehabilitation treatment. We can replace the 'Whole-brain Feature Extraction' module in the framework with feasible methods. Some specific implementations are given in Section II-F.

Two compared pipelines are designed for comparison. To make all the patients' injury areas in the same hemisphere, the conventional contra-ipsilesional mirroring routine is adopted on the whole-brain features, labeled as MCI Pipeline. In RIA Pipeline, we add a feature selection procedure to remedy the asymmetry problem, as shown in the box of Fig 1. Specifically, the whole-brain features are extracted from healthy controls with the same feature extraction procedure of patients. Then paired tests are used to check whether the mean of left and right lateral observations in healthy controls was equal, paired T-test for characteristics that hold the assumption of normality (determined by Shapiro's test, P > 0.05), otherwise paired Wilcoxon's test. This procedure is performed for every extracted feature and thus generates a p-value for each paired test. All significant attributes (P < 0.05, FDR corrected) are asymmetrical, while those remaining are regarded as left-right symmetry indices. Accordingly,

we apply the above indices to extracted features of patients to select symmetrical features, denoting the connectivity whose features are supposed to be left-right hemispheric balanced for healthy controls. Finally, features are sent to the classification and prediction models.

F. Preprocessing and Feature Extraction

The SPM-based package DPARSFA [27] was used to perform the standard resting-state BOLD (rsBOLD) signal preprocessing steps. Including the first 10 out of 240-time points were discarded, the spatial resolution was set at 2 mm^3, realigning, normalizing on Montreal Neurological Institute space with 3 mm isotropic voxel resampling, smoothing with a Gaussian smoothing kernel with 4-mm full width at half maximum, removing the linear trend of time courses and nuisance covariates with global signal regression and finally temporally filtering with 0.01-0.08 Hz. The preprocessing of the DTI signal was performed by using FSL tools (http://www.fmrib.ox.ac.uk/fsl/index.html) according to the following steps: skull stripping by function "bet2", eddy current and motion correction by function "EddyCorrect" and tensor fitting by function 'DTIFIT'.

Functional and structural features based on rsBOLD and DTI, respectively, were extracted by network construction based on a brain atlas and were carried out using in-house software. The functional features are some graph theoretic attributes extracted from rsBOLD. After preprocessing of rsBOLD, we constructed a fully weighted functional inter-connectivity network and intra-connectivity networks. The inter-connectivity network was obtained by using the Automated Anatomical Labeling (AAL) [28] atlas to define the nodes. Average all voxels in each node and then calculate the Pearson correlation coefficients between every pair of nodes as the connectivity. There are 116 brain regions in AAL, with odd numbers for the left hemisphere and even numbers for the right. For example, L1 and R2 are the same regions of the left and right hemispheres, respectively. Further, we constructed an intra-connectivity network for each brain region, in total 116 intra-connectivity networks. For each intra-connectivity network, the nodes were all the voxels corresponding to the brain region, and the Pearson correlation coefficient between each pair of voxels was the connectivity. The brain-connectivity toolbox [17] was utilized to extract a few attributes of the complex network according to graph theory, including local efficiency (LE), weighted degree (WD), betweenness centrality (BC), clustering coefficient (CC), global efficiency (GE), modularity (MO), characteristic path length (PL), small-worldness (SWN) [29]. For attributes of inter-connectivity and intra-connectivity, denoted here as .inter, .intra respectively. Overall, there are 10 attributes of BOLD: LE.inter, WD.inter, BC.inter, BC.intra, CC.inter, CC.intra, GE.intra, MO.intra, PL.intra, SWN.intra. The structural features of DTI are fractional anisotropy (FA) and mean diffusivity (MD). After preprocessing, we registered the AAL atlas to the native space of the reconstructed DTI. 'Dipy', a Python package was used to apply fiber tracking. Finally, FA and MD values of each brain region (nodes of the structural network) and fiber bundle counts between two regions (edges
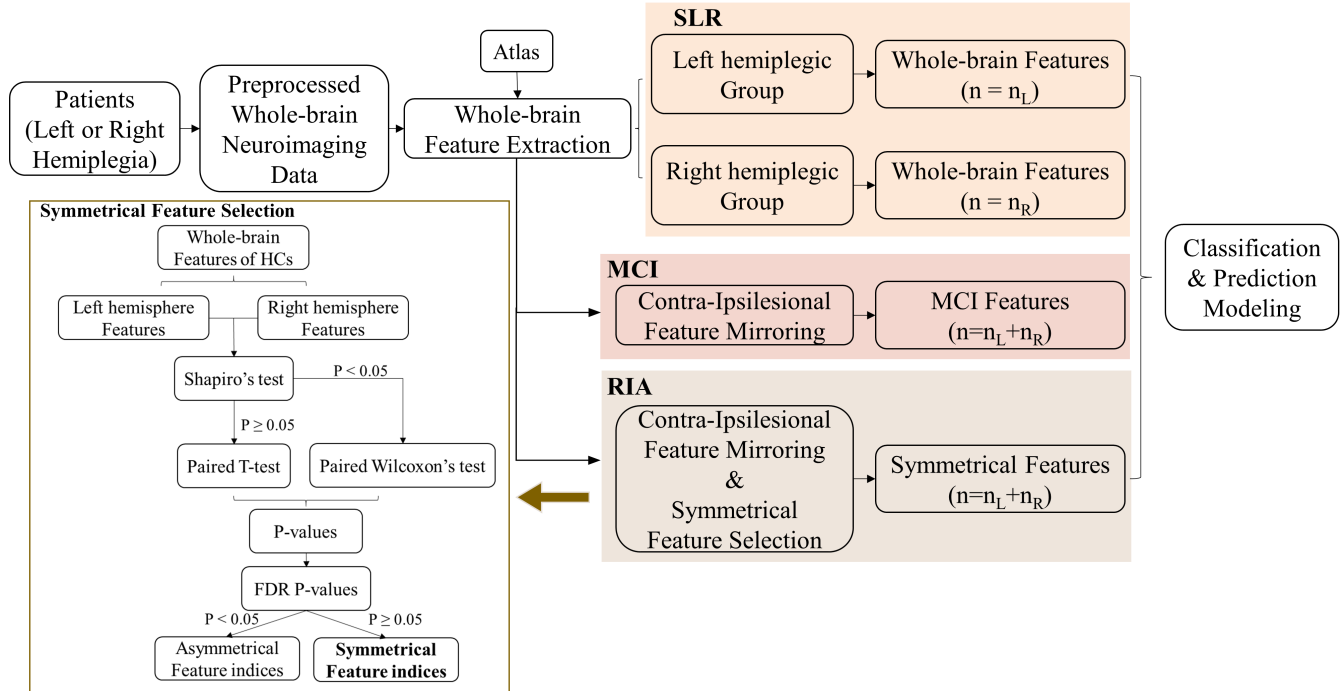
Pipeline

Fig. 1. The framework of proposed SLR modeling Pipeline (the top row), MCI (the middle row), and RIA (the bottom row) instead of SLR in the pipeline for comparison with the Contra-Ipsilesional hemispheres Mirrored and Symmetrical Feature Selection (shown in the left-bottom box). 'n$_L$' and 'n$_R$' denote the number of left/right hemiplegia patients, respectively. Symmetrical features represent that the features of the left and right hemispheric are symmetrical in healthy controls.
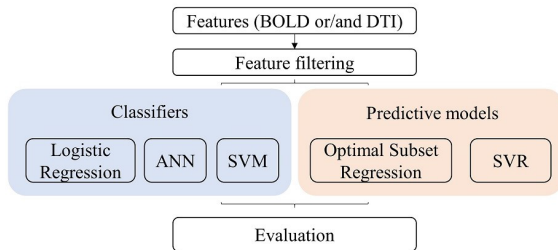

Fig. 2. Framework of Classifiers and Predictive models. The dashed line indicates the method used for the experiments shown in this paper.

of structural networks) were calculated. The feature vector for each subject is in total 1392 dimensions ($10 \times 116 = 1160$ BOLD features and $2 \times 116 = 232$ DTI features).

G. Classification and Prediction Modeling

Fig. 2 illustrates the detailed procedure of the "Classification and Prediction Modeling" module. The features obtained by the upstream of pipelines represent a way of brain imaging measurement and can be applied to explore underlying differences in physiological functions at the individual level. We then classify stroke patients into significant or non-significant responders to treatment and predict postinterventional outcomes by combining the machine learning framework and the above-calculated whole-brain features. It is worth noting that there is a mismatch between data size (dozens) and feature dimensionality (thousands). Feature

filtering is carried out to reduce the dimensionality of the data to avoid over-fitting. Subsequently, machine learning models are performed to classify and predict prognostic results, and evaluate features' importance. Brain regions contributing to the classification and prediction are then subjected to decoding analysis for impairment and recovery. Popular classifiers and predictive models are available for this framework [13], [15], [30], [31], [32], such as linear regression and its variants, artificial neural network (ANN) [33], support vector machine (SVM) [34] and support vector regression (SVR) [35].

Leave-one-out cross-validation (LOOCV) [36] is applied to estimate the performance of classifiers and predictive models. In the training process, we leave one sample as the test sample and utilize the remaining data for training. Specifically, Two-sample tests (T-test or Wilcoxon's test, uncorrected) were conducted for feature filtering in training data, across significant responders and non-significant responders on each attribute to identify the features that make a difference. Some machine learning frameworks were implemented in the experiment, logistic regression (LR), ANN with one hidden layer, SVM for classification, optimal subset regression (OSR) [37] derived from linear regression, and SVR for prediction. Besides the neurophysiological variables, initial clinical scores are significant predictors of outcomes in stroke patients [15], [38]. Therefore, we include the pre-interventional clinical scores into the predictors regarding prediction. The trained models are then employed to classify/predict the left-out sample.

TABLE II
LOOCV CLASSIFICATION ACCURACY OF ALL PIPELINES AND DATASETS

	Methods	BOLD	DTI	BOLD+DTI
LR	SLR Pipeline	**1.000** **L 1.000 / R 1.000**	**0.850** **L 0.833 / R 0.864**	**1.000** **L 1.000 / R 1.000**
	MCI Pipeline	0.875	0.750	0.875
	RIA Pipeline	0.875	0.675	0.875
ANN	SLR Pipeline	**0.875** **L 0.833 / R 0.909**	**0.75** **L 0.722 / R 0.773**	**0.925** **L 0.833 / R 1.000**
	MCI Pipeline	0.775	0.675	0.725
	RIA Pipeline	0.725	0.675	0.700
SVM	SLR Pipeline	**0.950** **L 0.944 / R 0.955**	**0.900** **L 0.833 / R 0.955**	**0.975** **L 1.000 / R 0.955**
	MCI Pipeline	0.900	0.800	0.925
	RIA Pipeline	0.925	0.675	0.900

L: left hemiplegia; R: right hemiplegia.

Repeat the operation until all patients are traversed to test the performances. We adapt accuracy, R^2 and mean absolute error (MAE) to evaluate models.

III. RESULTS

A. Asymmetrical and Significant Features

For all 40 healthy controls from HCP, BOLD and DTI features were extracted with the same pipeline of patients, and 258 of the total 1392 features were judged as asymmetrical, accounting for 18.8%. Fig. 3 shows the heatmap of left-right hemispheric asymmetry. For detailed information, see Supplementary Table S3.

We performed statistical tests which are the same as feature filtering in each pipeline but across all subjects and listed the significant features selected in Supplementary Table S4. The results show that 26 of 1160 BOLD features and 21 of 232 DTI features were significantly different for the left hemiplegic group, while the results were 41 BOLD features and 20 DTI features for the right hemiplegic group, more significant than the number of characteristics in MCI Pipeline and RIA Pipeline. Only a tiny proportion differed among thousands of features, contributing to the distinction between significant and non-significant responders. Separate modeling of left and right hemiplegic patients helps to detect more potentially significant features.

B. Performance on Classification

Table II lists the LOOCV accuracy of each pipeline/dataset for different methods. Besides, an exact number of correctly classified samples and the total sample amount for SLR Pipeline left and right models are summed to form the overall accuracy. The results of ANN are unstable due to the training parameter settings. To ensure reproducible and comparative results, we set seeds and fixed the best parameter in MCI Pipeline as the hidden layer's size. Different seeds and the hidden layer size did not affect the conclusions presented.

The results show that SLR Pipeline gives a 5-15% improvement to all datasets and methods. The sensitivity and specificity of each implementation are listed in Supplementary Table S5 and Supplementary Table S6 and the SLR Pipeline
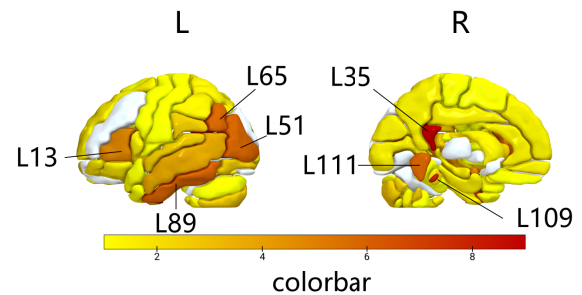

Fig. 3. Heatmap of Asymmetry in healthy controls. The color bar indicates the number of times the brain region was judged to be asymmetrical in 12 attributes.

significantly improves the classification performance was verified therefore. LR and SVM performed well when modeling left and right hemiplegia separately, and the classification accuracy/sensitivity/specificity is close to 1. As a remedy to asymmetry, RIA Pipeline does not help in the case of classification. For all implementations in our experiment, except BOLD attributes as the dataset in SVM, removing all asymmetrical features led to the same or worse LOOCV accuracy, resulting in increased sensitivity and decreased specificity, making it easier to determine patients as significant responders to reduce the possibility of missed diagnoses.

Using BOLD attributes alone as the dataset for all methods and pipelines gave better results than using DTI attributes alone. Across all machine learning methods, SLR Pipeline maintained and improved the accuracy when combining the two types of attributes. In contrast, most others have led to worse results than using BOLD alone. It implied that the BOLD and DTI partially overlap or contradict each other, and SLR Pipeline helps us discover complementary characteristics.

C. Performance on Prediction

The results in Table III show the R^2 and MAE for each implementation. The OSR method had the largest R^2 because it optimizes for sparsity, which is more conducive to finding strong predictors of the dependent variable and avoiding the phenomenon of overfitting. The dataset with a combination of BOLD and DTI features improved the effectiveness of the

TABLE III
LOOCV R2 AND MAE OF ALL PIPELINES AND INPUTS (INCLUDING PRE FMA-TOTAL)

	Methods	BOLD	DTI	BOLD+DTI
OSR	SLR Pipeline	**0.8556 (5.15)**	**0.8919 (4.70)**	**0.9466 (3.30)**
		L 0.7455 (7.52)	**L 0.8895 (4.66)**	**L 0.9410 (3.49)**
		R 0.9427 (3.21)	**R 0.8968 (4.74)**	**R 0.9523 (3.14)**
	MCI Pipeline	0.7601 (6.43)	0.7026 (7.44)	0.7820 (6.14)
	RIA Pipeline	0.7601 (6.43)	0.6643 (8.01)	0.7850 (5.99)
SVR	SLR Pipeline	0.5917 (9.27)	**0.7062 (7.53)**	0.6137 (8.83)
		L 0.4914 (10.94)	**L 0.4832 (10.18)**	L 0.4433 (10.73)
		R 0.6673 (7.90)	**R 0.8744 (5.37)**	R 0.7500 (7.28)
	MCI Pipeline	0.6118 (8.65)	0.5849 (8.89)	0.6007 (8.14)
	RIA Pipeline	**0.6301 (8.26)**	0.6773 (7.28)	**0.6816 (7.54)**

L: left hemiplegia; R: right hemiplegia. MAE in brackets

TABLE IV
NAMES OF INTERESTING REGIONS CORRESPONDING TO AAL-LABELS

Label	Abbrev.	Region name
L3, R4	SFGdor	Superial frontal gyrus, dorsolateral
L25, R26	ORBsupmed	Superior frontal gyrus, medial orbital
L35, R36	PCG	Posterior cingulate gyrus
L41, R42	AMYG	Amygdala
L43, R44	CAL	Calcarine fissure and surrounding cortex
L51, R52	MOG	middle occipital gyrus
L55, R56	FFG	Fusiform gyrus
L59, R60	SPG	Superior parietal gyrus
L61, R62	IPL	Inferior parietal, but supramarginal and angular gyri
L63, R64	SMG	Supramaginal gyrus
L67, R68	PCUN	Precuneus
L73, R74	PUT	Lenticular nucleus putamen
L87, R88	TPOmid	Temporal pole: middle temporal gyrus
L89, R90	ITG	Inferior temporal gyrus
L93, R94	CRBL_C2	Cerebellum_Crus2
L101, R102	CRBL7b	Cerebellum 7b
L113, R114	Vermis	Vermis 7, Vermis 8

OSR method compared to using them separately as inputs. This is because the number of predictors of OSR was limited to 5, and new invalid features were not selected, ensuring that the results were no less effective than the situation where only some features were used. SLR Pipeline improved the R^2 from approximately 0.75 to at least 0.85. Furthermore, the MAE is consistent with R^2. The larger the R^2, the smaller the MAE. The proposed SLR pipeline brings about a 40% reduction in MAE, indicating the reliability of the results. Similar to the case of classification, RIA does not improve the performance of the OSR predictive model. See Supplementary Fig.S2-S4 for specific feature selection, and there were precisely 5 features after feature filtering for RIA Pipeline with DTI as input, so no more selection was performed.

For the SVR model, SLR Pipeline was slightly more effective than MCI Pipeline. SVR aims to find a hyperplane that minimizes the distance from all the data without imposing too many restrictions on the features. It is often thought that dimensionality is a blessing for finding a support vector, but the rule failed in this experiment. The model became better with the asymmetry features removed. That is because only dozens of subjects were included, and almost all of them were selected as support vectors in the high-dimensional case. RIA can be regarded as a manual feature selection process to reduce dimensionality and assist the SVR in selecting more appropriate support vectors, thus improving prediction results. The model performed best when using only the DTI dataset in SLR Pipeline. This suggests that DTI features in prediction can be used as prognostic biomarkers, agreeing with previous findings [10], [39].

D. Importance of Predictors

Linear regression and its variants (LR and OSR in this experiment) autotune the parameters and select some of the features as predictors. Which brain regions chosen by the predictors are the most important for neurorehabilitation prognosis prediction? As illustrated in Fig. 4, by permutating explanatory variables, the order of importance of the predictors for the BOLD + DTI dataset represents the corresponding brain regions and their features. Box plots were overlapped to the bars to provide the distribution of the importance of the measure across 50 permutations. The names of interesting brain regions corresponding to the label of AAL-atlas and their abbreviations are shown in Table IV. Only a few features in the dataset were strong predictors of post-intervention FMA-total. Significant predictors in this experiment, including Pre FMA-total and a few specific functional brain regions, are consistent with previous studies.
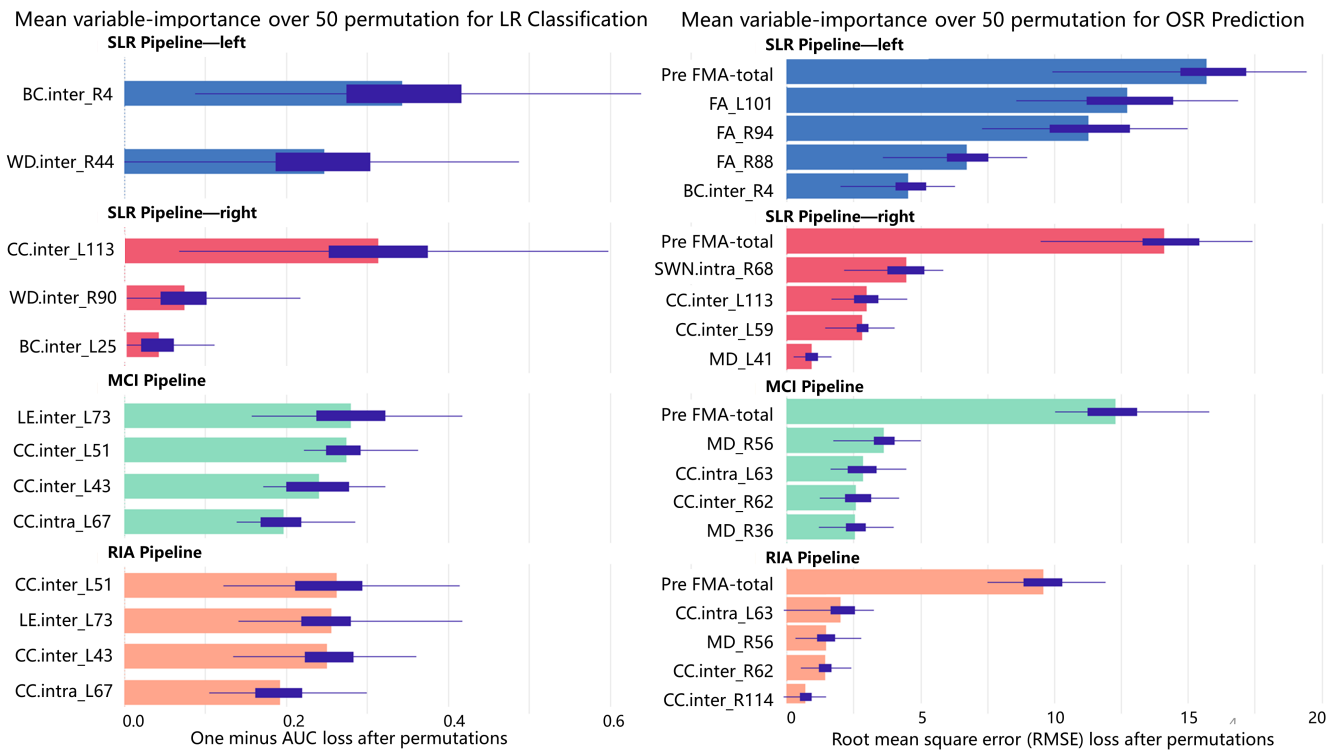
Fig. 4. Importance of predictors in classification for LR method and predictive models for OSR method with BOLD+DTI as input. Slight color bars represent the average prediction, whereas dark boxes represent the standard deviation over 50 permutations.

IV. DISCUSSION

Many studies have discussed left/right brain asymmetry from various aspects. Friedrich et al. [40] has found that we can identify left and right hemispheres in their low- and high-dimensional representation. Ortiz et al. [41] studied motor imagery performance after asymmetrical transcranial direct current stimulation. Obermaier et al. [42] introduced an HMM-based classifier allowing asymmetrical structures to help design the EEG-based BCI training paradigm. Due to the known asymmetries during motor imagery of right- and left-hand movement, the classifier of imagery tasks demonstrates an improvement of 9% for the classification accuracy. Unique patterns and brain changes in stroke recovery have been discovered in left- and right-hemispheric strokes [43], [44], [45], indicating that prognosis is significantly correlated with injured lateral. The human brain asymmetry is also validated in this study, via the changes in results brought by compared pipeline RIA (sometimes enhancing and sometimes worsening). Further, these changes indicate that the intrinsic asymmetry would confound the evidence for making decisions of classification and prediction. And the classification and prediction results will be significantly improved when separating the hemiplegic side for modeling. These facts would be helpful for revealing possible hemispheric recovery mechanisms in response to brain damage and offering guidance for stimulation during treatment, to the end of which, more and further exploration is deserved.

We utilized BOLD and DTI features to perform the pipelines since they provide functional and structural information, respectively. FA and MD as DTI features, BC and CC as BOLD features have been validated to be significantly

different before and after therapy in previous studies [39], [46]. These four features might be used as biomarkers to predict post-intervention outcomes for stroke patients, although their performance varies and the importance thereof is still unclear. In this study, we evaluated the role of each data set and found that our proposed SLR pipeline worked consistently across all datasets to enhance performance. The results show that BOLD provides more recovery information in classification, while DTI has more potential to predict post-intervention scores. However, DTI has a large amount of structural asymmetry (in healthy controls), and in RIA, only 4 features were left to be used as discriminators, with reduced reliability. In general, the combination of BOLD and DTI improves the prediction results, while only slight changes occur to the classification. For each pipeline, BOLD + DTI does not always perform better than BOLD or DTI alone, and there may exist a contradictory relationship between the two kinds of features, suggesting that the graphical properties of the BOLD network contain a large amount of structural information represented by DTI.

Noticeably, across all classification and prediction methods employed in this experiment, the traditional statistical method (LR and OSR) performed better than algorithmic modeling (ANN, SVM, and SVR), with higher accuracy and super interpretability, mainly due to the sample size [47]. For detailed functions of specific brain regions in classification, the chosen features remained the same after RIA, but their importance changed–the clustering coefficient of L51 (MOG.L) elevated, while others remained unchanged. The most important two regions, L51 and L73 (PUT.L), are associated with cognitive and motor function [48], [49]. Among the subjects of

this experiment, L73/R74 was with the highest frequency of injury, and the LR classifiers of MCI and RIA indicated that the higher the local efficiency of L73, the more likely the patient was to be a responder of treatment. For SLR Pipeline, regions associated with visual (R4, R44, L25) and learning (R90) [50], [51] appeared essential. Generally, these regions are more often associated with learning cognitive functions. We conjecture that BCI training requires patients to engage in motor imagery tasks and that a corresponding neural circuit is to control the training process. For prediction, the pre-intervention FMA-total is always the most critical factor and other brain regions vary for pipelines.

There are some limitations to this study. Brain atlases provide an objective definition of nodes and thus are usually used to construct brain networks. Node definitions will change the graph's topology and influence the functional relationships between nodes and hence affecting our results. Differences in atlas-related metrics for resting-state networks [52] and structural networks [53] have been found. In the present study, we only used the AAL atlas to define nodes, and future studies will focus on selecting brain atlases and fusing them to construct multilevel networks.

Another drawback of this study was feature filtering. We adopted the LR method to perform feature selection automatically while no selection for ANN and the result that LR has higher accuracy than that of ANN reveals the importance of feature filtering to some extent. In the pipeline, we performed an uncorrected t-test/Wilcoxon as preliminary filtration to detect as many potentially significant features as possible, leading to the relatively high dimensionality of the feature space. But in other ways, it may introduce redundant features and cause pitfalls in the classifier training. In the future, the t-test/Wilcoxon-test filter can be upgraded to a more well-designed and robust method. Finally, the data size of corresponding patients is relatively small and as the left and right hemiplegia are distinguished beforehand, the proposed SLR modeling pipeline is suitable for unilateral stroke patients, and how it performs on patients with multiple brain lesions remains unknown.

V. Conclusion

For more accuracy of neuro-rehabilitation prognosis prediction applied in the clinic, a left and right hemiplegia separated modeling pipeline SLR is initiated in our study to eliminate the asymmetry, especially in brain structure, to improve the performance of classification and prediction. BOLD and DTI are both used for validating the classification and prediction accuracy of recovery levels after BCI-assisted training of stroke patients with upper limb dysfunction. The comparative experiment is designed to compare the classification and prediction accuracy provided by three modeling pipelines, i.e., SLR, MCI, and RIA, with the same three classification and two prediction methods. The LOOCV results show that SLR can efficiently distinguish the significant and non-significant responders and make more accurate predictions of post-intervention FMA-total. The overall effect of targeted modeling for left and right hemiplegic patients is better than MCI and RIA. This is the first time rs-fMRI and DTI are applied separately on the left and right hemiplegic patients to evaluate the rehabilitation outcomes. The final classification and prediction models obtained from SLR could be used to help clinicians make a more reliable rehabilitation program for stroke patients.

References

[1] R. von Bernhardi, L. Eugenín-von Bernhardi, and J. Eugenn, "What is neural plasticity?" in *The Plastic Brain* (Advances in Experimental Medicine and Biology), R. von Bernhardi, J. Eugenn, and K. J. Müller, Eds. Cham, Switzerland: Springer, 2017, pp. 1–15.

[2] P. M. Rossini and S. Rossi, "Transcranial magnetic stimulation: Diagnostic, therapeutic, and research potential," *Neurology*, vol. 68, no. 7, pp. 484–488, Feb. 2007.

[3] J. J. Daly and J. R. Wolpaw, "Brain-computer interfaces in neurological rehabilitation," *Lancet Neurol.*, vol. 7, no. 11, pp. 1032–1043, 2008.

[4] W.-C. Huang, Y.-J. Chen, C.-L. Chien, H. Kashima, and K.-C. Lin, "Constraint-induced movement therapy as a paradigm of translational research in neurorehabilitation: Reviews and prospects," *Amer. J. Transl. Res.*, vol. 3, no. 1, pp. 48–60, 2010.

[5] A. M. Barrett, M. Oh-Park, P. Chen, and N. L. Ifejika, "Neurorehabilitation," *Neurol., Clin. Pract.*, vol. 3, no. 6, pp. 484–492, 2013.

[6] M. H. Lee, C. D. Smyser, and J. S. Shimony, "Resting-state fMRI: A review of methods and clinical applications," *Amer. J. Neuroradiol.*, vol. 34, no. 10, pp. 1866–1872, Oct. 2013.

[7] B. Bodini and O. Ciccarelli, "Diffusion MRI in neurological disorders," in *Diffusion MRI*, H. Johansen-Berg and T. E. J. Behrens, Eds., 2nd ed. San Diego, CA, USA: Academic, 2014, ch. 11, pp. 241–255.

[8] E. R. Buch, S. Rizk, P. Nicolo, L. G. Cohen, A. Schnider, and A. G. Guggisberg, "Predicting motor improvement after stroke with clinical assessment and diffusion tensor imaging," *Neurology*, vol. 86, no. 20, pp. 1924–1925, May 2016.

[9] A. Kuceyeski et al., "Structural connectome disruption at baseline predicts 6-months post-stroke outcome," *Hum. Brain Mapping*, vol. 37, no. 7, pp. 2587–2601, Jul. 2016.

[10] L. M. Moura et al., "Diffusion tensor imaging biomarkers to predict motor outcomes in stroke: A narrative review," *Frontiers Neurol.*, vol. 10, May 2019, Art. no. 445.

[11] E. Tavazzi, N. Bergsland, A. Pirastru, M. Cazzoli, V. Blasi, and F. Baglio, "MRI markers of functional connectivity and tissue microstructure in stroke-related motor rehabilitation: A systematic review," *NeuroImage, Clin.*, vol. 33, Jan. 2022, Art. no. 102931.

[12] I. R. König et al., "Predicting long-term outcome after acute ischemic stroke: A simple index works in patients from controlled clinical trials," *Stroke*, vol. 39, no. 6, pp. 1821–1826, Jun. 2008.

[13] C. M. Stinear, W. D. Byblow, S. J. Ackerley, M.-C. Smith, V. M. Borges, and P. A. Barber, "PREP2: A biomarker-based algorithm for predicting upper limb function after stroke," *Ann. Clin. Transl. Neurol.*, vol. 4, no. 11, pp. 811–820, Nov. 2017.

[14] J. Puig et al., "Resting-state functional connectivity magnetic resonance imaging and outcome after acute stroke," *Stroke*, vol. 49, no. 10, pp. 2353–2360, 2018.

[15] C. Tozlu et al., "Machine learning methods predict individual upper-limb motor impairment following therapy in chronic stroke," *Neurorehabilitation Neural Repair*, vol. 34, no. 5, pp. 428–439, May 2020.

[16] E. Bullmore and O. Sporns, "Complex brain networks: Graph theoretical analysis of structural and functional systems," *Nature Rev. Neurosci.*, vol. 10, no. 3, pp. 186–198, Mar. 2009.

[17] M. Rubinov and O. Sporns, "Complex network measures of brain connectivity: Uses and interpretations," *NeuroImage*, vol. 52, no. 3, pp. 1059–1069, Sep. 2010.

[18] M. Khosla, K. Jamison, G. H. Ngo, A. Kuceyeski, and M. R. Sabuncu, "Machine learning in resting-state fMRI analysis," *Magn. Reson. Imag.*, vol. 64, pp. 101–121, Dec. 2019.

[19] X.-Z. Kong et al., "Mapping cortical brain asymmetry in 17,141 healthy individuals worldwide via the ENIGMA consortium," *Proc. Nat. Acad. Sci. USA*, vol. 115, no. 22, pp. E5154–E5163, 2018.

[20] P. Bisiacchi and E. Cainelli, "Structural and functional brain asymmetries in the early phases of life: A scoping review," *Brain Struct. Function*, vol. 227, no. 2, pp. 479–496, Mar. 2022.

[21] S. Naghdi, N. N. Ansari, K. Mansouri, and S. Hasson, "A neurophysiological and clinical study of Brunnstrom recovery stages in the upper limb following stroke," *Brain Injury*, vol. 24, no. 11, pp. 1372–1378, Oct. 2010.

[22] A. A. Puzi, S. N. Sidek, H. M. Rosly, N. Daud, and H. M. Yusof, "Modified Ashworth scale (MAS) model based on clinical data measurement towards quantitative evaluation of upper limb spasticity," *IOP Conf. Ser., Mater. Sci. Eng.*, vol. 260, Nov. 2017, Art. no. 012024.

[23] S. Y. Bookheimer et al., "The lifespan human connectome project in aging: An overview," *NeuroImage*, vol. 185, pp. 335–348, Oct. 2019.

[24] Q. Wu et al., "Brain functional networks study of subacute stroke patients with upper limb dysfunction after comprehensive rehabilitation including BCI training," *Frontiers Neurol.*, vol. 10, Jan. 2020, Art. no. 1419.

[25] D. J. Gladstone, C. J. Danells, and S. E. Black, "The Fugl-Meyer assessment of motor recovery after stroke: A critical review of its measurement properties," *Neurorehabilitation Neural Repair*, vol. 16, no. 3, pp. 232–240, Sep. 2002.

[26] J. van der Lee, H. Beckerman, G. Lankhorst, and L. Bouter, "The responsiveness of the action research arm test and the Fugl-Meyer assessment scale in chronic stroke patients," *J. Rehabil. Med., Off. J. UEMS Eur. Board Phys. Rehabil. Med.*, vol. 33, no. 3, pp. 110–113, 2001.

[27] C.-G. Yan, and Y.-F. Zang, "DPARSF: A MATLAB toolbox for 'pipeline' data analysis of resting-state fMRI," *Frontiers Syst. Neurosci.*, vol. 4, May 2010, Art. no. 13.

[28] N. Tzourio-Mazoyer et al., "Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain," *NeuroImage*, vol. 15, no. 1, pp. 273–289, Jan. 2002.

[29] M. D. Humphries and K. Gurney, "Network 'small-world-ness': A quantitative method for determining canonical network equivalence," *PLoS ONE*, vol. 3, no. 4, Apr. 2008, Art. no. e0002051.

[30] C. M. Stinear, P. A. Barber, M. Petoe, S. Anwar, and W. D. Byblow, "The PREP algorithm predicts potential for upper limb recovery after stroke," *Brain, J. Neurol.*, vol. 135, no. 8, pp. 2527–2535, Aug. 2012.

[31] W.-Y. Lin et al., "Predicting post-stroke activities of daily living through a machine learning-based approach on initiating rehabilitation," *Int. J. Med. Informat.*, vol. 111, pp. 159–164, Mar. 2018.

[32] P. Sale et al., "Predicting motor and cognitive improvement through machine learning algorithm in human subject that underwent a rehabilitation treatment in the early stage of stroke," *J. Stroke Cerebrovascular Diseases*, vol. 27, no. 11, pp. 2962–2972, Nov. 2018.

[33] A. K. Jain, J. Mao, and K. M. Mohiuddin, "Artificial neural networks: A tutorial," *Computer*, vol. 29, no. 3, pp. 31–44, Mar. 1996.

[34] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, Jan. 1995.

[35] H. Drucker, C. J. C. Burges, L. Kaufman, A. Smola, and V. Vapnik, "Support vector regression machines," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 9, 1996, pp. 155–161.

[36] J. Lever, M. Krzywinski, and N. Altman, "Model selection and overfitting," *Nature Methods*, vol. 13, no. 9, pp. 703–704, Sep. 2016.

[37] D. Boyce, A. Farhi, and R. Weischedel, *Optimal Subset Selection: Multiple Regression, Interdependence and Optimal Network Algorithms*. Berlin, Germany: Springer, 2013.

[38] J. Y. Jo et al., "Prediction of motor recovery using quantitative parameters of motor evoked potential in patients with stroke," *Ann. Rehabil. Med.*, vol. 40, no. 5, pp. 806–815, 2016.

[39] D.-M. Wang, J. Li, J.-R. Liu, and H.-Y. Hu, "Diffusion tensor imaging predicts long-term motor functional outcome in patients with acute supratentorial intracranial hemorrhage," *Cerebrovascular Diseases*, vol. 34, no. 3, pp. 199–205, 2012.

[40] P. Friedrich et al., "Is it left or is it right? A classification approach for investigating hemispheric differences in low and high dimensionality," *Brain Struct. Function*, vol. 227, no. 2, pp. 425–440, Mar. 2022.

[41] M. Ortiz, E. Iáñez, J. A. Gaxiola-Tirado, D. Gutiérrez, and J. M. Azorín, "Study of the functional brain connectivity and lower-limb motor imagery performance after transcranial direct current stimulation," *Int. J. Neural Syst.*, vol. 30, no. 8, Aug. 2020, Art. no. 2050038.

[42] B. Obermaier, C. Munteanu, A. Rosa, and G. Pfurtscheller, "Asymmetric hemisphere modeling in an offline brain-computer interface," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 31, no. 4, pp. 536–540, 2001.

[43] Y. Chen et al., "Common and unique structural plasticity after left and right hemisphere stroke," *J. Cerebral Blood Flow Metabolism*, vol. 41, no. 12, pp. 3350–3364, Dec. 2021.

[44] R. B. S. C. Garbus, J. M. Prado-Rico, A. G. Nardini, and S. M. S. F. Freitas, "Immediate effects of arm reaching training in standing on postural control differ between right and left stroke individuals," *J. Stroke Cerebrovascular Diseases*, vol. 30, no. 9, Sep. 2021, Art. no. 105984.

[45] C. Rovito, A. Park, R. Goldstein, R. Zafonte, R. Black-Schaffer, and J. C. Schneider, "A retrospective cohort comparing left and right middle cerebral artery ischemic stroke functional outcomes in acute inpatient rehabilitation," *PM&R*, vol. 13, no. 7, pp. 666–673, Jul. 2021.

[46] M. Mazrooyisebdani et al., "Evaluation of changes in the motor network following BCI therapy based on graph theory analysis," *Frontiers Neurosci.*, vol. 12, Nov. 2018, Art. no. 861.

[47] B. Efron, "Prediction, estimation, and attribution," *J. Amer. Stat. Assoc.*, vol. 115, no. 530, pp. 636–655, Apr. 2020.

[48] M. Packard and B. Knowlton, "Learning and memory functions of the basal ganglia," *Annu. Rev. Neurosci.*, vol. 25, pp. 563–593, Mar. 2002.

[49] L. A. Renier, I. Anurova, A. G. De Volder, S. Carlson, J. Van Meter, and J. P. Rauschecker, "Preserved functional specialization for spatial processing in the middle occipital gyrus of the early blind," *Neuron*, vol. 68, no. 1, pp. 138–148, Oct. 2010.

[50] S. H. Creem and D. R. Proffitt, "Defining the cortical visual systems: 'What', 'where', and 'how,'" *Acta Psychol.*, vol. 107, no. 1, pp. 43–68, 2001.

[51] I. I. Goldberg, M. Harel, and R. Malach, "When the brain loses its self: Prefrontal inactivation during sensorimotor processing," *Neuron*, vol. 50, no. 2, pp. 329–339, Apr. 2006.

[52] J. Wang et al., "Parcellation-dependent small-world brain functional networks: A resting-state fMRI study," *Hum. Brain Mapping*, vol. 30, no. 5, pp. 1511–1523, May 2009.

[53] A. Zalesky et al., "Whole-brain anatomical networks: Does the choice of nodes matter?" *NeuroImage*, vol. 50, no. 3, pp. 970–983, Apr. 2010.