scientific data

OPEN

DATA DESCRIPTOR

A multi-paradigm EEG dataset for studying upper limb rehabilitation exercises

Wenwen Chang[1] ✉, Weixuan Kong[1], Guanghui Yan[1], Renjie Lv[1], Kaiyue Du[1], Muhammad Tariq Sadiq[2], Bin Guo[3], Rong Yin[4] & Xuan Liu[4]

Most stroke survivors experience persistent upper limb motor dysfunction, and brain-computer interface (BCI) rehabilitation technologies have been widely explored to address this issue. However, systematic comparisons and analyses of differences among rehabilitation paradigms remain challenging due to the lack of multi-paradigm EEG datasets from the same subjects. This study aims to construct an EEG dataset that collects various rehabilitation paradigms for the same subjects. A total of 28 healthy subjects were recruited, and EEG data were collected under six types of upper limb rehabilitation paradigms. Each paradigm involves two or three actions, including grasping and releasing with the left, right, or both hands. The dataset includes both raw EEG signals and preprocessed versions with bandpass filtering and artifact removal. This resource will support studies comparing the neural mechanisms underlying different rehabilitation paradigms and aid in the development of optimized rehabilitation strategies.

Background & Summary

Stroke, a cerebrovascular disease characterized by the sudden onset of rapidly developing focal or diffuse neurological deficits, leaves over 80% of survivors with residual motor dysfunction in their limbs[1,2]. Among these, the loss of upper limb function is particularly critical, as most activities of daily living require the use of the upper limbs. This significantly diminishes patients' quality of life and their independence in daily activities.

As a novel human-computer interaction technology, brain-computer interfaces (BCIs) have garnered significant attention in rehabilitation research in recent years and have demonstrated substantial potential[3–5]. BCI systems primarily rely on two categories of neurophysiological signals: electrophysiological signals and hemodynamic signals. Electrophysiological activity is recorded and analyzed through modalities such as electroencephalography (EEG), electrocorticography (ECoG), and magnetoencephalography (MEG)[6], while hemodynamic signals are measured using functional magnetic resonance imaging (fMRI)[7] and functional near-infrared spectroscopy (fNIRS)[8]. Among these, EEG has emerged as the most widely adopted technique for monitoring brain activity in BCI systems due to its advantages of non-invasiveness, low cost, and operational simplicity[9–11].

Among the various rehabilitation training modalities, motor imagery (MI) is one of the most widely utilized techniques in neurorehabilitation[12,13]. When a subject imagines a limb movement, event-related desynchronization (ERD) or event-related synchronization (ERS) occurs in the sensorimotor cortex. These EEG features can decode movement intentions and facilitate neural circuit reorganization in patients[14,15]. Furthermore, a variety of upper limb rehabilitation approaches—including mirror therapy (MT)[16,17], virtual reality (VR)[18,19], soft rehabilitation gloves[20–22], and mirror-assisted glove rehabilitation[23], all demonstrate distinct neural signatures through EEG signal analysis.

Researchers have extensively explored brain-computer interface (BCI)-based neurorehabilitation methods. However, existing datasets for upper limb functional training predominantly focus on motor imagery (MI) paradigms[24–26]. While some datasets incorporate additional paradigms, such as motor execution (ME)[27,28] and cross-session motor imagery analyses[29], However, there is a lack of datasets that collect multiple types of

[1]School of Electronic and Information Engineering, Lanzhou Jiaotong University, Lanzhou, 730070, China. [2]School of Computer Science and Electronic Engineering, University of Essex, Colchester Campus, Colchester, CO4 3SQ, UK. [3]School of Computer Science, Northwestern Polytechnical University, Xi 'an, 710129, China. [4]Gansu Provincial Maternity and Child-care Hospital (Gansu Provincial Central Hospital), Lanzhou, 730079, China. ✉e-mail: changww2013@126.com

movement data across different upper limb rehabilitation paradigms for the same subject. Therefore, this paper proposes a dataset that includes multiple experimental paradigms for the same subject to compare the characteristics and advantages of different rehabilitation paradigms.

Upper limb rehabilitation typically begins with gross motor movements, which are crucial during early recovery from paralysis[30]. This dataset focuses on the collection of EEG signals during hand grasping actions. Although such actions do not involve proximal movements like those of the elbow joint, they represent a fundamental mode of fine motor control and are directly related to the core needs of daily functional rehabilitation in stroke patients. Research indicates that hand grasping ability is a prerequisite for performing activities of daily living, and that even the unaffected hand of chronic stroke patients exhibits deficits in fine motor skills[31,32]. In terms of rehabilitation applications, grasping actions directly correspond to the most critical functional demands in daily life (e.g., holding objects, eating), and their training intensity is significantly lower than that of multi-joint movements, making them suitable for BCI training in patients[33].

This dataset is also suitable for research on low-frequency oscillations (LFOs). LFOs plays a critical role in brain functional network information transmission and integration, with growing research confirming its importance for post-stroke motor recovery[30,34]. The raw EEG data in this study contains sufficient low-frequency components, which can serve as a comparative reference for studying LFOs-related post-stroke hand rehabilitation.

In the dataset presented in this study, we collected data from 28 healthy subjects across six distinct upper limb rehabilitation paradigms. Subjects performed experimental tasks in response to visual and auditory cues in a quiet environment. All blocks were continuously monitored by the experimenters to ensure data richness and statistical power.

Following data collection, we validated the presence of ERD/ERS phenomena across two categories of MI tasks, confirming the dataset's reliability and neurophysiological plausibility. Furthermore, we conducted machine learning-based classification analyses on EEG data from all paradigms, with experimental results demonstrating the dataset's robust classifiability. Finally, by constructing functional brain networks for the six rehabilitation paradigms, we found differences in network topology diagrams among some paradigms, providing novel train of thought for optimizing rehabilitation protocols. Compared to existing datasets, the proposed dataset integrates multiple upper limb rehabilitation paradigms and offers resources for studying the impact of various upper limb rehabilitation methods on brain neural networks.

Methods

Subjects and experimental procedure. The dataset comprises data from 28 healthy subjects (14 males, 14 females). Aged between 20 and 33, all right-handed. During the recruitment of volunteers, the purpose of the experiment was clearly communicated. Subjects were screened to meet the following criteria: no history of neurological or psychiatric disorders, normal cognitive function, normal vision, and adherence to pre-experiment requirements, which included sufficient rest (no sleep deprivation the night before) and abstaining from caffeinated beverages (e.g., coffee, tea) on the day of the experiment. Informed consent was obtained from all volunteers for data collection, and the data were anonymized prior to publication. The study was approved by the Ethics Committee of Gansu Provincial Maternity and Child-care Hospital (Gansu Provincial Central Hospital)((2023) GSFY Ethics [90]) and conducted in strict accordance with the Declaration of Helsinki. All participants provided written informed consent, confirming their understanding of the experimental procedures and agreement to both participate in the study and allow anonymous data publication.

The experimental procedure is illustrated in Fig. 1. After arriving at the laboratory, subjects were first familiarized with the experimental environment to ensure psychological relaxation. To further support subjects' mental relaxation and reduce potential stress-related artifacts, the experimental environment was designed to be quiet and isolated, with soft ambient lighting. Before the experiment began, each subject had a free conversation period with the experimenter. During this time, the experimenter provided background information, and the subjects had the opportunity to ask questions about anything they were unsure of, helping them adapt to the environment and feel relaxed. Additionally, after each experiment, a brief informal assessment of the subjects' mental and physical state was conducted through short verbal interviews. Rest periods were flexibly adjusted based on their feedback to maintain comfort and sustained attention throughout the experiment. These procedures were designed to minimize mental tension in the subjects, as mental tension can introduce muscle and ocular artifacts into the EEG data. Reducing such tension helps improve the overall reliability of the signal.

Subsequently, the EEG cap was put on the subjects, and conductive gel was applied to ensure proper electrode contact. ECG and EOG electrodes were attached. Subjects were then briefed on the experimental requirements, including that they were not allowed to speak, avoid extraneous body movements, and sustain focused attention throughout each trial. Before the start of each paradigm, subjects performed a small number of practice trials to ensure their understanding of the task. The testing system and EEG signal quality were confirmed to be normal, after which the formal experiment commenced. During the experiment, the subjects were seated 60 centimeters in front of the computer screen, with their eyes level with the screen.

The stimulus scenes used in the experiment were created using U3D (Unity 3D), with all scenes rendered at a refresh rate of 60 frames per second (fps). The stimulus categories are shown in Fig. 1 and consist of six components. The fist-opening/closing animations for both hands were modeled in MAYA. During export, the animation frame rate was synchronized with the U3D scene refresh rate (60 fps) to ensure smooth playback. Motion synchronization latency between the mirror glove and soft rehabilitation glove was maintained at $\leq$1 second. Two soft rehabilitation gloves provided assisted motion, with each assisted palm extension achieving an amplitude of $\geq$190° and assisted fist closure limited to $\leq$0°.

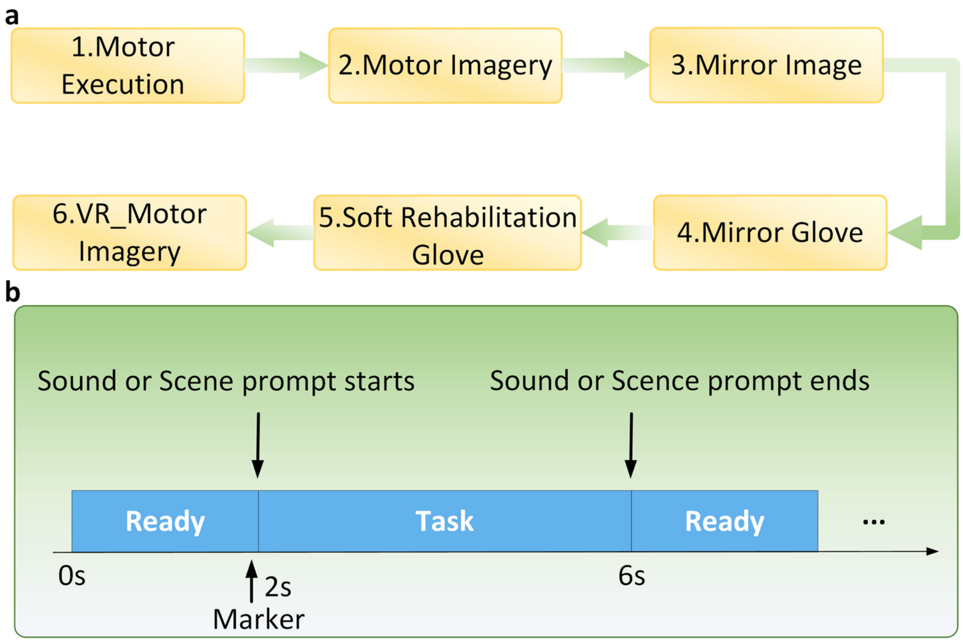

Fig. 1 Experimental procedures. The order of data acquisition (**a**), one trial (**b**).

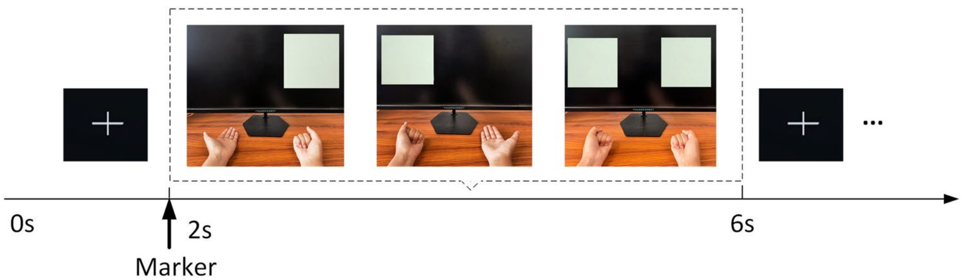

Fig. 2 Experimental scene of motion execution.

Experimental paradigms. This experiment includes six different rehabilitation paradigms. The specific design of each paradigm is as follows:

Paradigm 1: Motor Execution

As shown in Fig. 2. Subjects place both arms forward on the table, focus on a cross "+" displayed on the screen. When the cross disappears, they observe the location of a square on the screen. If the square appears on the left, the subject performs a grasp-and-release movement with the left hand; if it appears on the right, the subject performs a grasp-and-release movement with the right hand; if squares appear on both sides, the subject performs a grasp-and-release movement with both hands.

Paradigm 2: Motor Imagery

As shown in Fig. 3. Subjects place both arms forward on the table, focus on a cross "+" displayed on the screen. When the cross disappears, they observe the location of a square on the screen. If the square appears on the left, the subject imagines performing a grasp-and-release movement with the left hand without actual movement; if on the right, the subject imagines the same with the right hand without actual movement; if squares appear on both sides, both hands are imagined to perform the movement without any actual execution.

Paradigm 3: VR_Motor Imagery

As shown in Fig. 4. Subjects place both arms forward on the table and wear VR glasses, focusing on a cross "+" within the VR display. When the cross disappears, subjects observe the movements of virtual hands. If the virtual left hand moves, they learn the movement and imagine their own left hand performing the grasp-and-release movement without actual movement; similarly, if the virtual right hand moves, they imagine the same with their right hand. If both virtual hands move, they imagine performing the movement with both hands without any actual movement.

For the above three paradigms, each trial lasts 4 seconds, followed by a 2-second rest. Each session lasts for 3 minutes, during which each hand performs 10 trials. The experiment is repeated 4 times, resulting in 40 trials per hand.

Paradigm 4: Mirror Therapy

Fig. 3 Experimental scene of motion imagery.

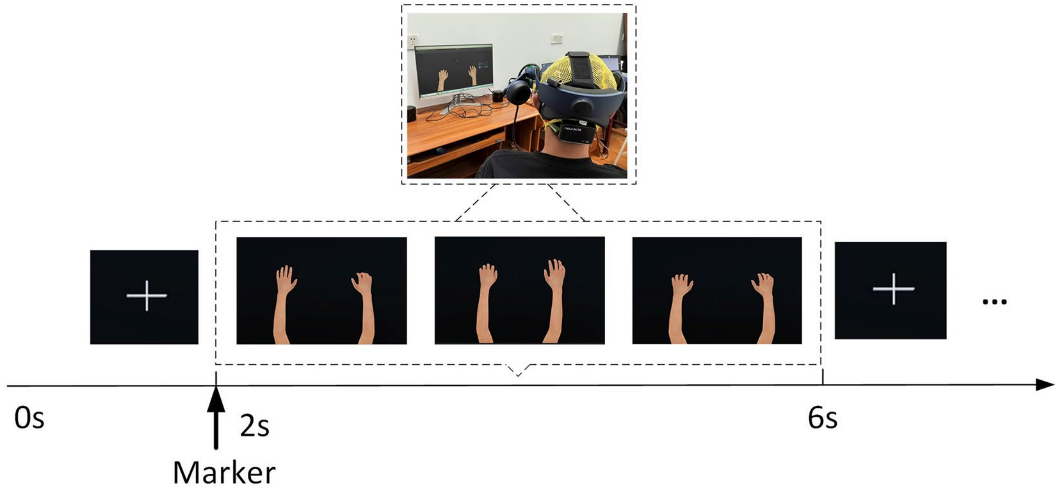

Fig. 4 Experimental scene of VR_motor imagery.

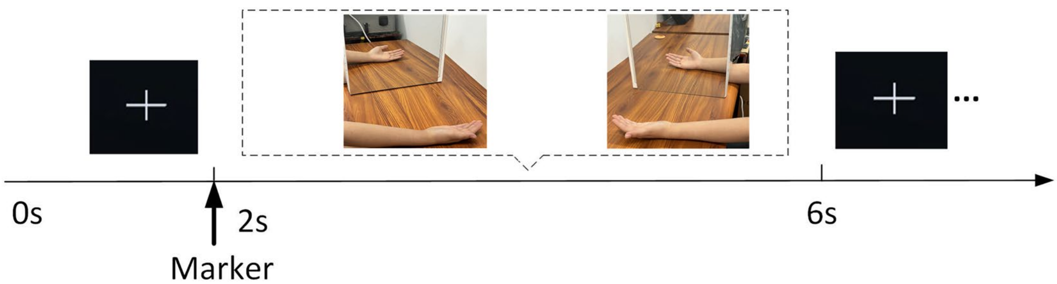

Fig. 5 Experimental scene of mirror therapy.

As shown in Fig. 5. One hand is placed inside a device box (simulating the affected side), and the other hand is placed in front of a mirror (simulating the healthy side). Upon hearing an instruction from the experimental system, subjects focus on the mirrored image of the hand. The "healthy side" hand performs a grasp-and-release movement, which is reflected in the mirror to simulate the movement of the "affected side." This aids the subject in imagining the movement of the "affected side." Each mirror therapy trial lasts 4 seconds, followed by a 2-second rest. Each session lasts 2 minutes, with 20 trials per hand. The experiment is repeated 4 times, totaling 40 mirror therapy trials per hand.

Paradigm 5: Soft Rehabilitation Gloves

As shown in Fig. 6. The subject placed their arms forward on the table, put on soft rehabilitation gloves, and focusing on the fixed cross "+" displayed on the screen. During the mission, they focusing on experiencing glove-driven gripping and release movements. There are three types of tasks: left-handed, right-handed, and two-handed conditions. In different tasks, only the designated hand is wearing soft rehabilitation gloves. For example, when the left hand assists movement, only the left hand wears gloves; when the right hand assists

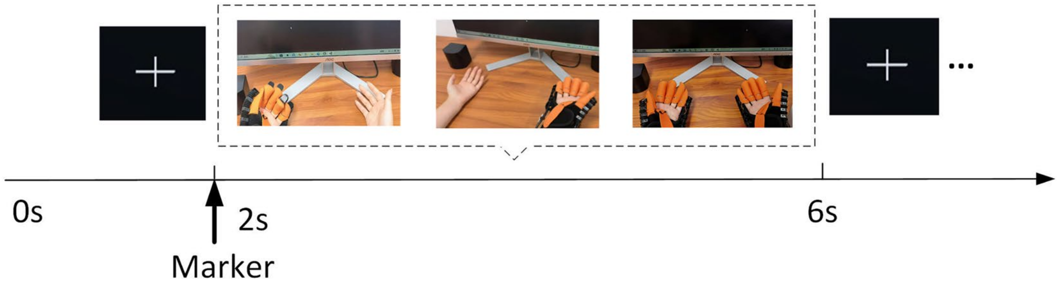

Fig. 6 Experimental scene of soft rehabilitation gloves.

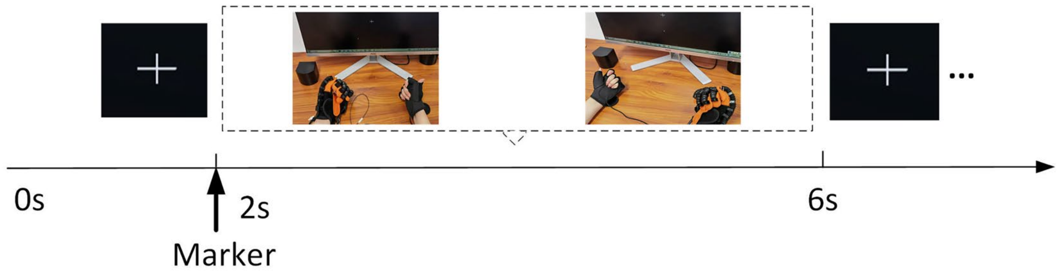

Fig. 7 Experimental scene of mirror glove.

movement, only the right hand wears it; in the case of both hands, both hands are gloved. Each trial lasts 4 seconds, followed by a 2-second rest. Each session lasts 2 minutes, with 20 trials per hand. The experiment is repeated 6 times, totaling 40 assisted movement trials per hand.

Paradigm 6: Mirror Glove

As shown in Fig. 7. One hand wears an assistive glove, while the other hand wears a mirrored glove. Both arms are placed forward on the table, focusing on a cross "+" displayed on the screen. Upon receiving a command from the experimental system, the hand with the mirrored glove performs a grasp-and-release movement that drives the other hand to perform the same movement. Each mirror-assisted session lasts 4 seconds, followed by a 2-second rest. Each session lasts 2 minutes, with 20 trials per hand. The experiment is repeated 4 times, totaling 40 mirror-assisted movement trials per hand.

Rest periods are adjusted between every two sessions based on the subject's condition.

Data acquisition. EEG data were acquired using the BRK-NSW 2.0 system (Neuracle Technology Co., Ltd., China), which comprises an EEG cap, a wireless EEG amplifier, a multi-parameter synchronizer, and EEG acquisition software. According to the BRK-NSW 2.0 specifications, the system provides high-precision EEG signal measurements and adopts a DC-coupled amplification mode. During the data acquisition process, it can retain all low-frequency signals, thereby achieving high-precision electroencephalographic signal measurement. The EEG cap's electrodes were positioned according to the international 10–20 system, with a total of 59 EEG channels and a sampling rate of 1000 Hz. As shown in Figs. 8 and 9. conductive gel was injected into each electrode channel before the experiment. The impedance of all electrodes was checked using the BRK-NSW 2.0 software system. Data collection commenced only after all electrode channels achieved an impedance of ≤20 kΩ, indicating proper preparation. Throughout the experiment, electrode impedance was intermittently rechecked to ensure all channels remained within the acceptable impedance range.

Data Records

This section describes the organization and storage structure of the EEG dataset, which is accessible through the Figshare data repository service.

The directory structure is illustrated in Fig. 10 and contains two types of data:

(1) Raw EEG data[35]: https://doi.org/10.6084/m9.figshare.28831730.v2; stored in folders named "S2XX," where "XX" represents the subject ID. Each subject's folder contains 26 subfolders, corresponding to the 26 experimental sessions. Each experimental session is saved in.set format.
(2) Preprocessed EEG data[36]: https://doi.org/10.6084/m9.figshare.28837016.v2; saved in the format "S2XX_preICA.set".(In addition, the preprocessed MATLAB code is included in the dataset and can be accessed directly).

Raw EEG dataset. All raw EEG data from the subjects are stored in the "Raw" directory. Each subject has a dedicated folder (labeled as S201,…, S206, S208,…, S217, S219, …, S230), within which data from 26

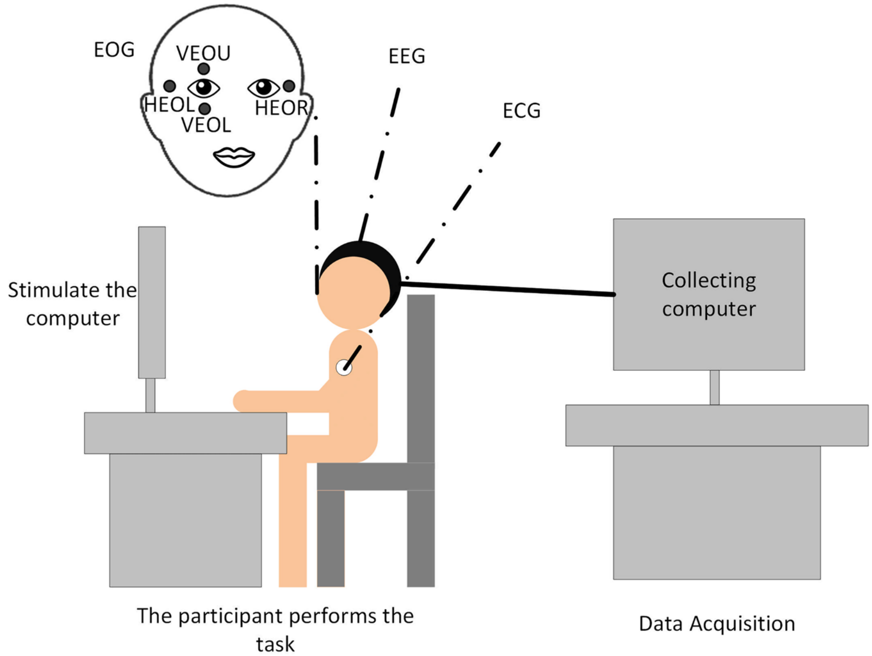

Fig. 8 The left panel shows the placement positions of electroencephalography (EEG), electrooculography (EOG), and electrocardiography (ECG) electrodes. EOG electrodes were positioned around the eyes: two placed above and below the left eye to detect vertical eye movements, and two placed laterally to both eyes to record horizontal eye movements. ECG electrodes were placed on the upper chest or near the clavicle to capture cardiac activity.

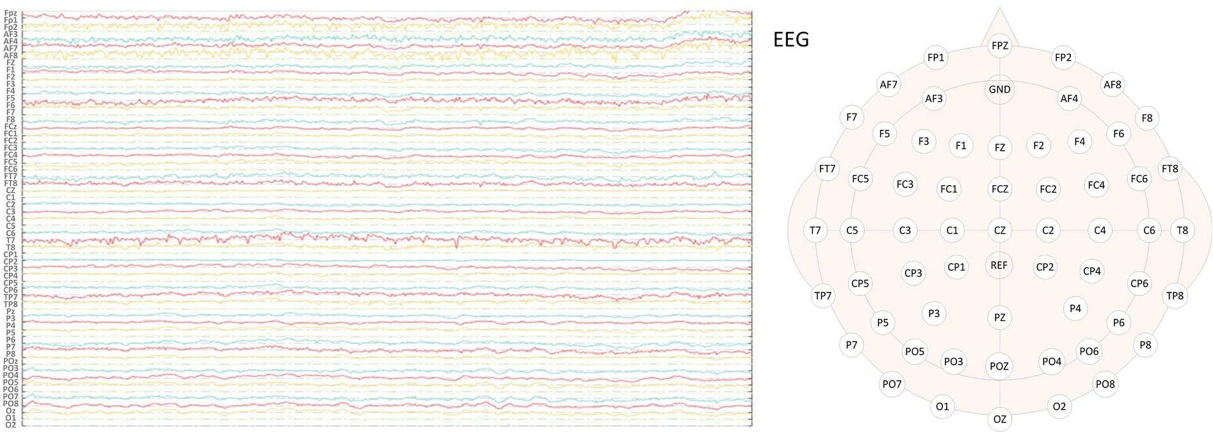

Fig. 9 The left side is the original EEG signal map, and the right side is the EEG cap electrode distribution location map.

experimental sessions are stored. These include 4 sessions of motor execution, 4 sessions of motor imagery guided by square positions on a computer screen, 4 sessions of mirror therapy, 4 sessions of mirror-gloves therapy, 4 sessions of VR-based motor imagery with virtual hand guidance, and 6 sessions of assisted movement using soft rehabilitation gloves.

The definitions and meanings of the event labels used in the dataset are provided in Table 1.

After loading the raw data using the EEGLAB toolbox, a structured variable named EEG is generated. The details in Fig. 10 list the names and corresponding meanings of the important fields in this structure. For information on all field names and their corresponding meanings, please refer to the "README" file in the dataset.

Preprocessed EEG dataset. After data collection, the collected data were preprocessed using the EEGLAB toolbox in MATLAB (R2023b). First, all behavioral samples from each subject were merged. The merged data were then subjected to downsampling, baseline correction, filtering, segmentation, and manual removal of bad segments. Following preprocessing, the original data, with a sampling rate of 1000 Hz per second, were downsampled to 512 Hz per second. The data were filtered within a frequency range of 0.1 Hz to 100 Hz. After

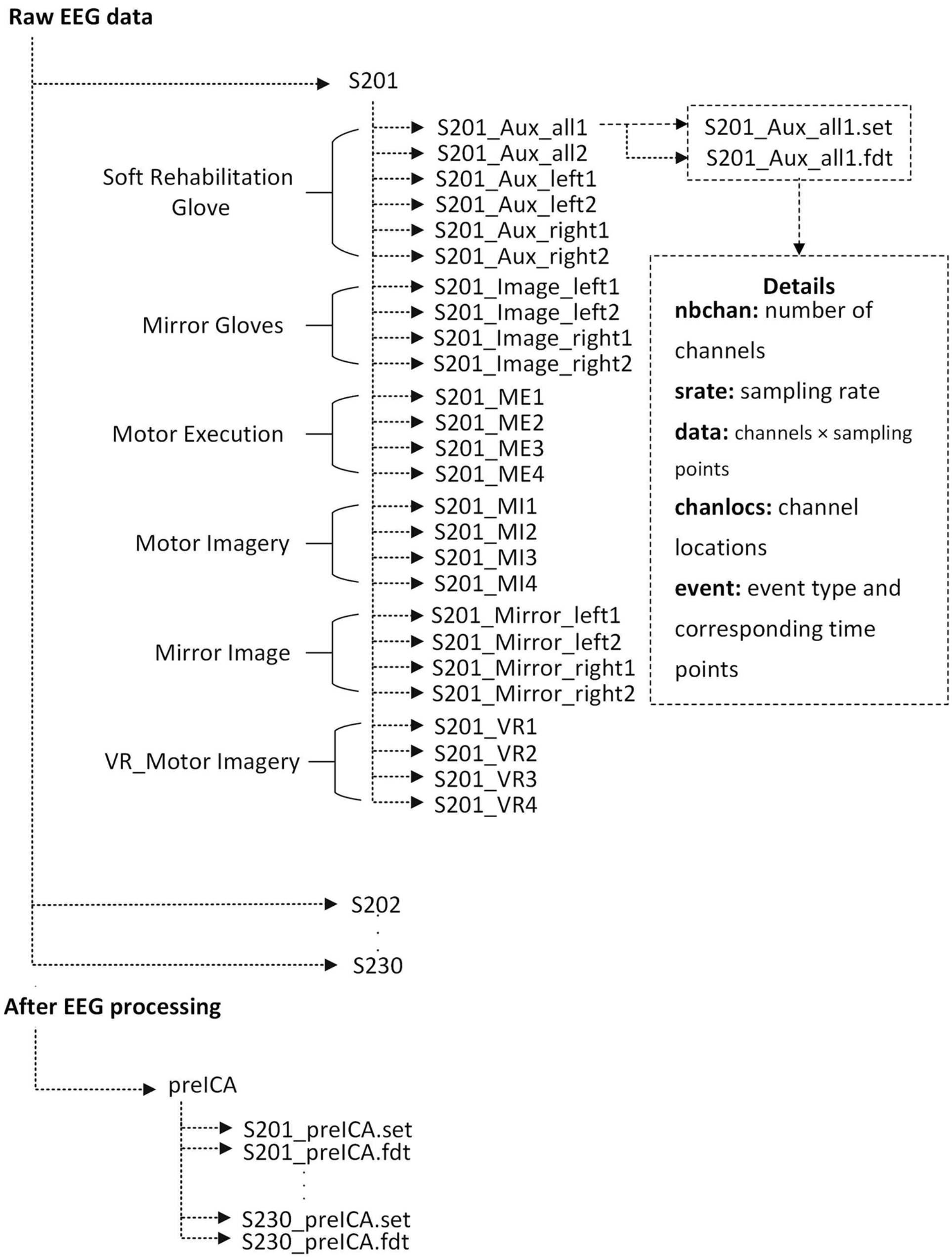

Fig. 10 The directory tree of the EEG file.

Event marker			
Event	**Left hand**	**Right hand**	**Both Hands**
Marker	1	2	3

Table 1. Event Marker Definitions.

preprocessing, the ".set" files contain EEG structure variables, with key fields such as: the data field formatted as (channels × time points × trials) and the event field containing event labels and their onset timestamps.

Technical Validation

To verify the validity of the dataset, focus on the following four aspects:

(1) Whether the EEG recordings are temporally synchronized with the event annotations.
(2) Whether event-related desynchronization/synchronization (ERD/ERS) phenomena can be observed during motor imagery tasks.

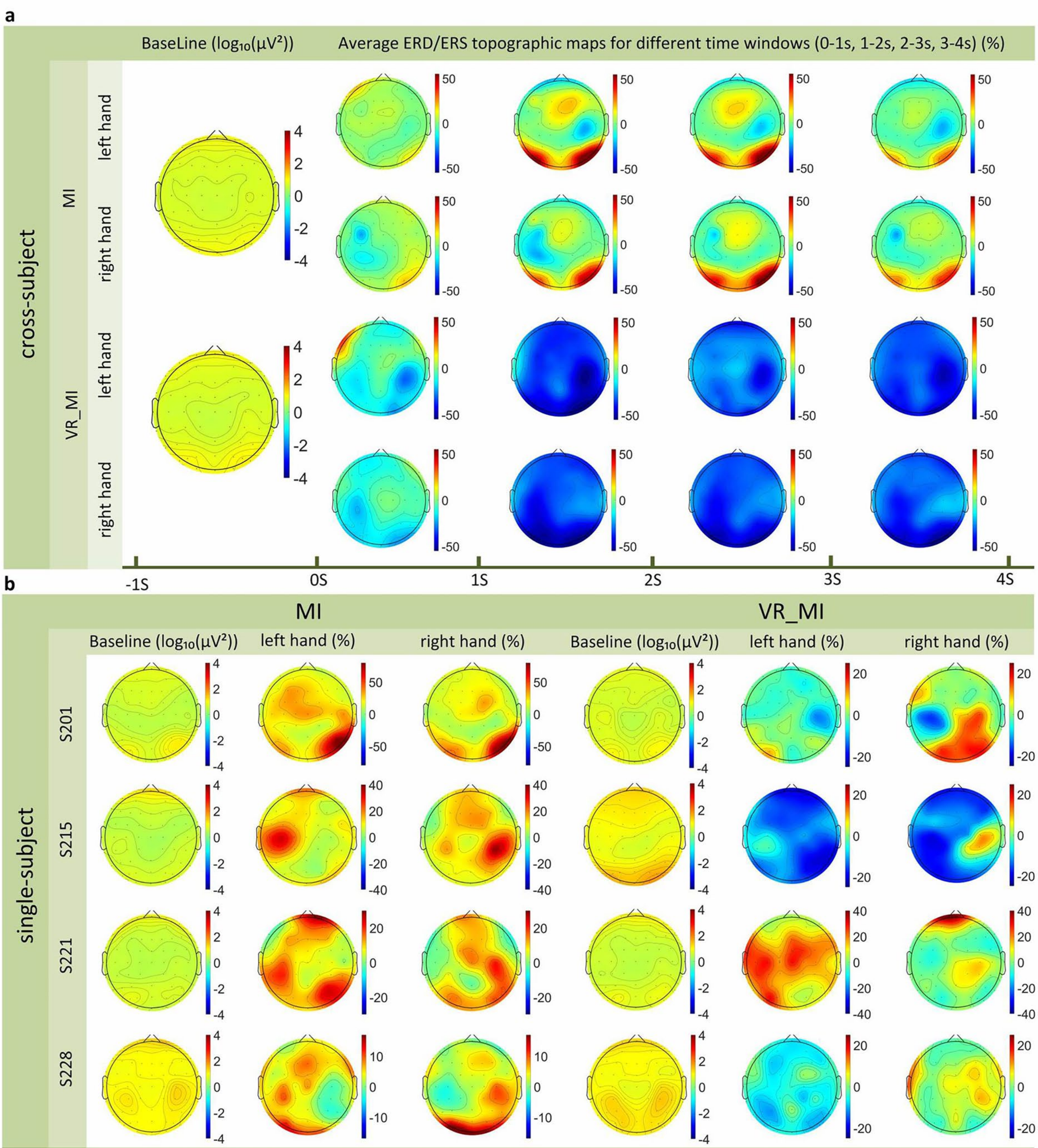

Fig. 11 Part a is the cross-subject analysis, with 0 s in the timeline being the time when the label marker is presented. Part b is the single-subject analysis, where the baseline power is expressed as $\log_{10}(\mu V^2)$, and the ERD/ERS is given as a percentage.

(3) The classification accuracy of different types within the same rehabilitation paradigm.
(4) The functional brain network patterns across different rehabilitation paradigms.

Synchronization between data and tasks. To ensure that event markers were temporally aligned with the onset of task-related EEG data, a U3D-based system was employed. Based on predefined conditions for each scenario, the system transmitted different event types via a serial port to the event buffer of the Neuracle data acquisition software. In paradigms such as motor execution and motor imagery, all experimental scenes were presented with a black background. When a task cue appeared, its onset time and type were simultaneously recorded within the Neuracle system. This synchronization between task events and data acquisition was achieved using a multi-parameter synchronizer provided by Neuracle. When the display computer detected a specific stimulus related to a motor paradigm, it triggered the synchronizer to send a timestamp signal to the data acquisition

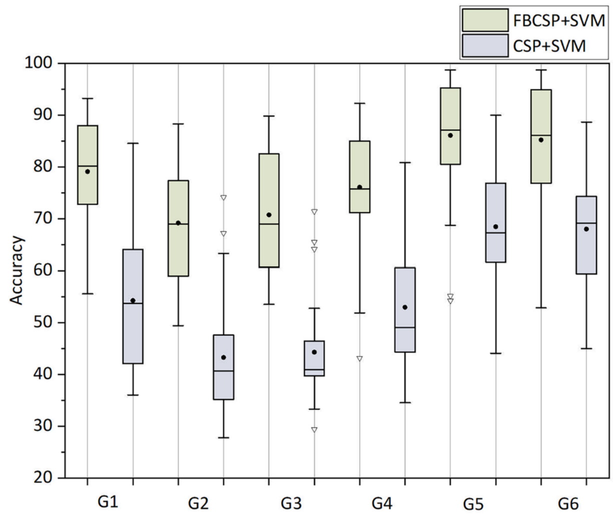

Fig. 12 Single-subject classification accuracy chart.

class	Soft rehabilitation glove	Motor execution	Motor imagery	VR_Motor imagery	Mirror glove	Mirror therapy
Accuracy	47.6	42.8	47.5	50.2	63.4	60.3

Table 2. Multi-subject classification accuracy table.

computer. All timestamps were stored in a.bdf file and used to align physiological recordings with experimental events, thereby ensuring synchronized data collection. It should be noted that in the paradigms involving mirror movement and mirror-assisted movement, auditory cues were used to prompt the subjects. As individual responses to auditory stimuli varied slightly, the event markers may have a minor deviation from the actual onset of the movement.

ERD/ERS. This section validates the effectiveness of two EEG datasets related to motor imagery tasks by calculating event-related desynchronization (ERD) and event-related synchronization (ERS). The study focuses on left-hand and right-hand grasping tasks, primarily analyzing EEG activity from 59 channels. The calculation method involves applying a bandpass filter (selecting the 8–13 Hz alpha band, as this band contains the mu rhythm closely associated with somatosensory and motor activities, thereby enabling more precise extraction of neural oscillatory activities related to sensorimotor functions) to the simply preprocessed data. Subsequently, continuous wavelet transform is applied to compute power values during the task period (i.e., the motor imagery phase after the cue appears) and the baseline period (i.e., the resting state before the task). The ERD/ERS percentage is then calculated according to the formula (task power - baseline power)/baseline power × 100%, and finally visualized intuitively via scalp topographical maps. For the baseline period, the power values of this segment were log-transformed and then converted into scalp topographic maps for visualization.

Two types of data were calculated. As shown in Fig. 11. The first type (cross-subject analysis): displays the average power topographic maps of the baseline period (-1-0 s) for all subjects, and also calculates the average ERD/ERS topographic maps for different task time windows (0-1 s, 1-2 s, 2-3 s, 3-4 s). This part aims to present the baseline activity and the dynamic evolution of neural activity along the task progression across subjects. The second type (single-subject analysis): we randomly selected four subjects (S201, S210, S215, S228), and displayed their respective baseline power topographic maps and corresponding ERD/ERS topographic maps (the baseline period here is uniformly -1-0s, and the task period is uniformly 0-2 s). This part aims to intuitively present the individual differences in baseline activity and task responses, as well as the general signal quality.

Classification accuracy. To assess whether the collected dataset can distinguish between different task states, we conducted a classification experiment on the EEG data of each subject after preprocessing. The methods used for classification include CSP + SVM and FBCSP + SVM. CSP (Common Spatial Pattern) is a classical method for feature extraction in EEG signals. The basic idea of the CSP method is as follows: (1) For each class of data, compute the covariance matrix and normalize it. Then, perform generalized eigenvalue decomposition on each class's covariance matrix to obtain spatial filtering matrices and eigenvalues. (2) Sort the eigenvalues in descending order, select the most significant eigenvectors, and apply a whitening transformation to ensure eigenvalue stability. (3) Apply the whitening transformation to the covariance matrix and perform feature decomposition to obtain the final CSP transformation matrix. (4) Finally, apply the CSP transformation matrix to each trial's data and calculate the log variance as features, thus extracting the most discriminative features. The FBCSP (Filter Bank CSP) method is an extension of the CSP method, which enhances the representational capacity of features by introducing frequency band division and filtering. Its basic principles are as follows: (1) Frequency

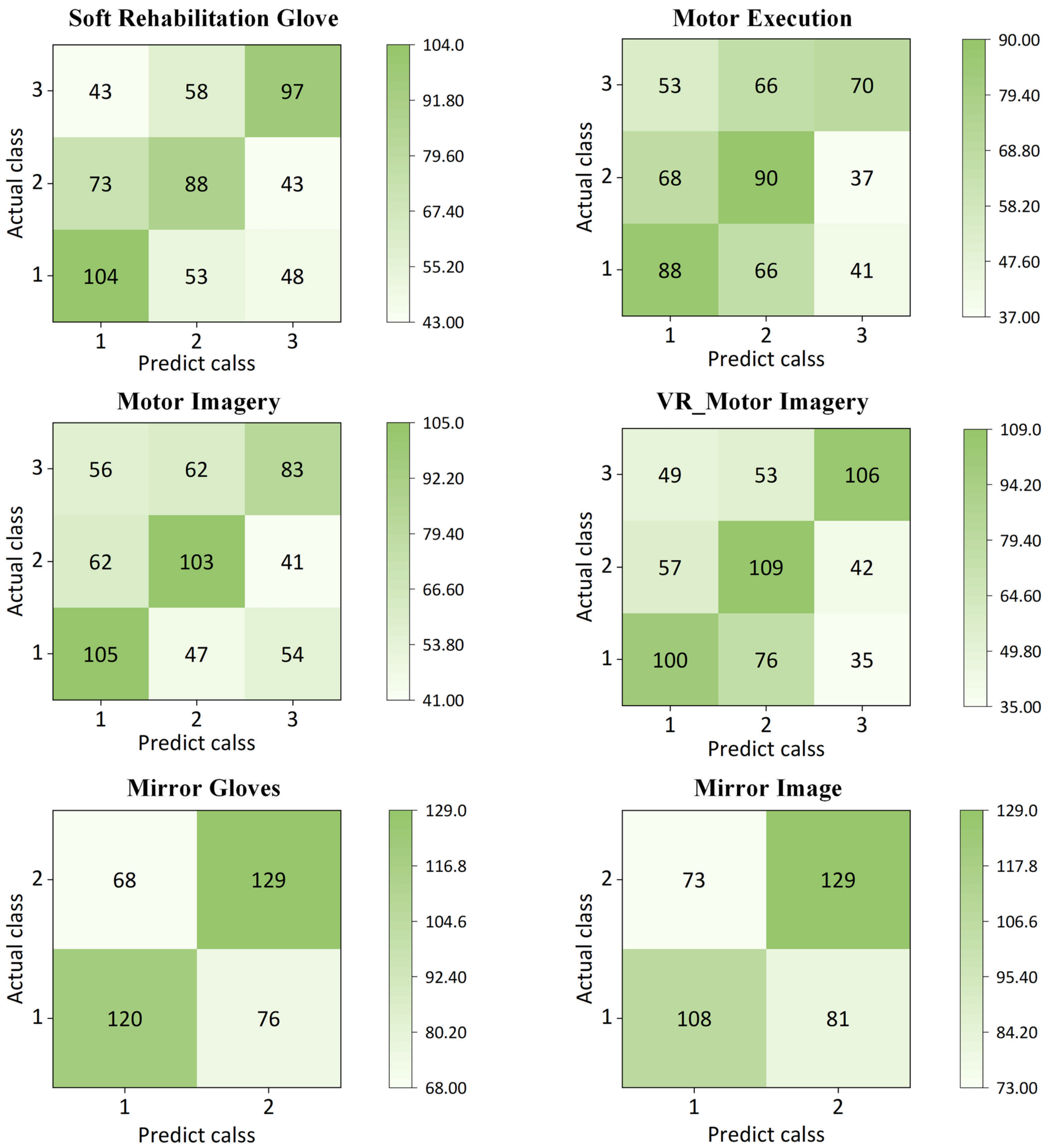

Fig. 13 Shows the confusion matrix of classification results for multiple subjects. The horizontal axis represents the predicted class labels, while the vertical axis represents the true class labels. Task labels are encoded as: 1 = left hand, 2 = right hand, 3 = both hands. The number in each cell indicates the count of samples where the predicted class matches the true class. Higher values along the anti-diagonal indicate higher classification accuracy.

band division: The raw EEG signal is first divided into multiple sub-bands. This study utilized six frequency bands covering 4–28 Hz (4–8 Hz, 8–12 Hz, 12–16 Hz, 16–20 Hz, 20–24 Hz, 24–28 Hz). (2) Sub-band feature extraction: The CSP method is applied to each sub-band to extract its features. (3) Feature selection: From the features generated by each frequency band, the top two and bottom two CSP features are selected, forming a 24-dimensional feature vector. Compared to the traditional CSP method, FBCSP decomposes the EEG signal into multiple frequency bands and applies CSP separately in each sub-band, ultimately producing a richer set of log-variance features, thereby improving classification accuracy and robustness.

Support Vector Machine (SVM) is then used to classify the generated feature vectors from the two classes. For each experiment, 10-fold cross-validation was performed on the dataset, and the average result was taken as the final classification accuracy.

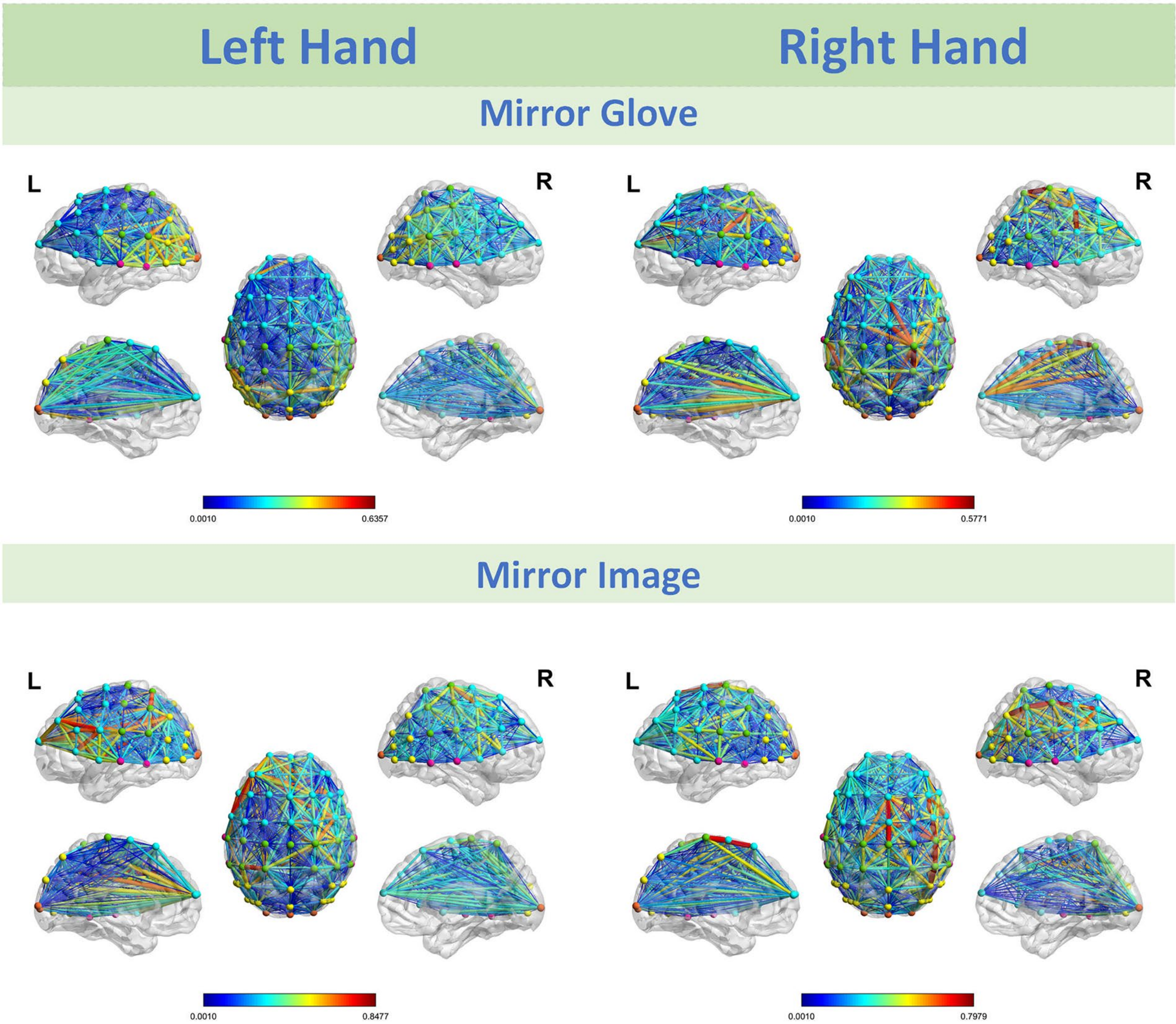

Fig. 14 Brain Functional Network Connectivity Diagrams 1.

Figure 12 presents the specific classification accuracy results (n = 28). In the CSP + SVM model, the average accuracies (mean ± standard deviation) for experiments G1 to G6 (G1: flexible rehabilitation glove, G2: motor execution, G3: motor imagery, G4: virtual reality motor imagery, G5: mirror glove, G6: mirror therapy) were 53.12% ± 13.22%, 42.37% ± 9.34%, 43.73% ± 8.25%, 53.37% ± 12.22%, 68.20% ± 11.35%, and 67.99% ± 10.87%, respectively. In comparison, the FBCSP + SVM model achieved higher accuracy across all tasks, with results of 79.11% ± 9.56%, 69.03% ± 10.71%, 70.77% ± 11.83%, 77.20% ± 9.63%, 86.59% ± 10.98%, and 85.88% ± 9.75%, respectively.

As shown in Table 2, the classification accuracy was validated across multiple subjects. The method used for classification was to apply the SVM to the feature vectors generated by FBCSP. Figure 13 presents the confusion matrix results for multiple subjects.

To evaluate the relative effectiveness of our dataset, we referenced another motor imagery EEG dataset involving healthy subjects (n = 52), which reported an average binary classification accuracy of 67.46% ± 13.17% using the CSP + FLDA method[37]. The performance of the CSP + SVM model used in our study for binary classification tasks was G5: 68.20% ± 11.35%; G6: 67.99% ± 10.87%, showing highly similar accuracy rates. Furthermore, for the three-class tasks (G1-G4), the accuracy ranged from 42.37% ± 9.34% to 53.37% ± 12.22%, which also yielded results consistent with our expectations: as the number of categories increases, the classification difficulty increases, and the accuracy correspondingly decreases. Additionally, we referenced a published EEG dataset from acute stroke patients related to motor imagery (n = 50). That study reported average binary classification accuracies of 55.57%, 57.57%, 61.20%, and 72.21% using the CSP + LDA, FBCSP + SVM, TSLDA + DGFDRM, and TWFB + DGFDM methods, respectively[24]. In our experiment, using the same method (FBCSP + SVM) on data from healthy subjects yielded higher classification accuracy (G5: 86.59% ± 10.98%, G6: 85.88% ± 9.75%). This difference was also expected, as acute stroke patients may have difficulty maintaining a sitting position and steadily performing motor imagery paradigms for extended periods.

In summary, the above demonstrates the reliability of our dataset.

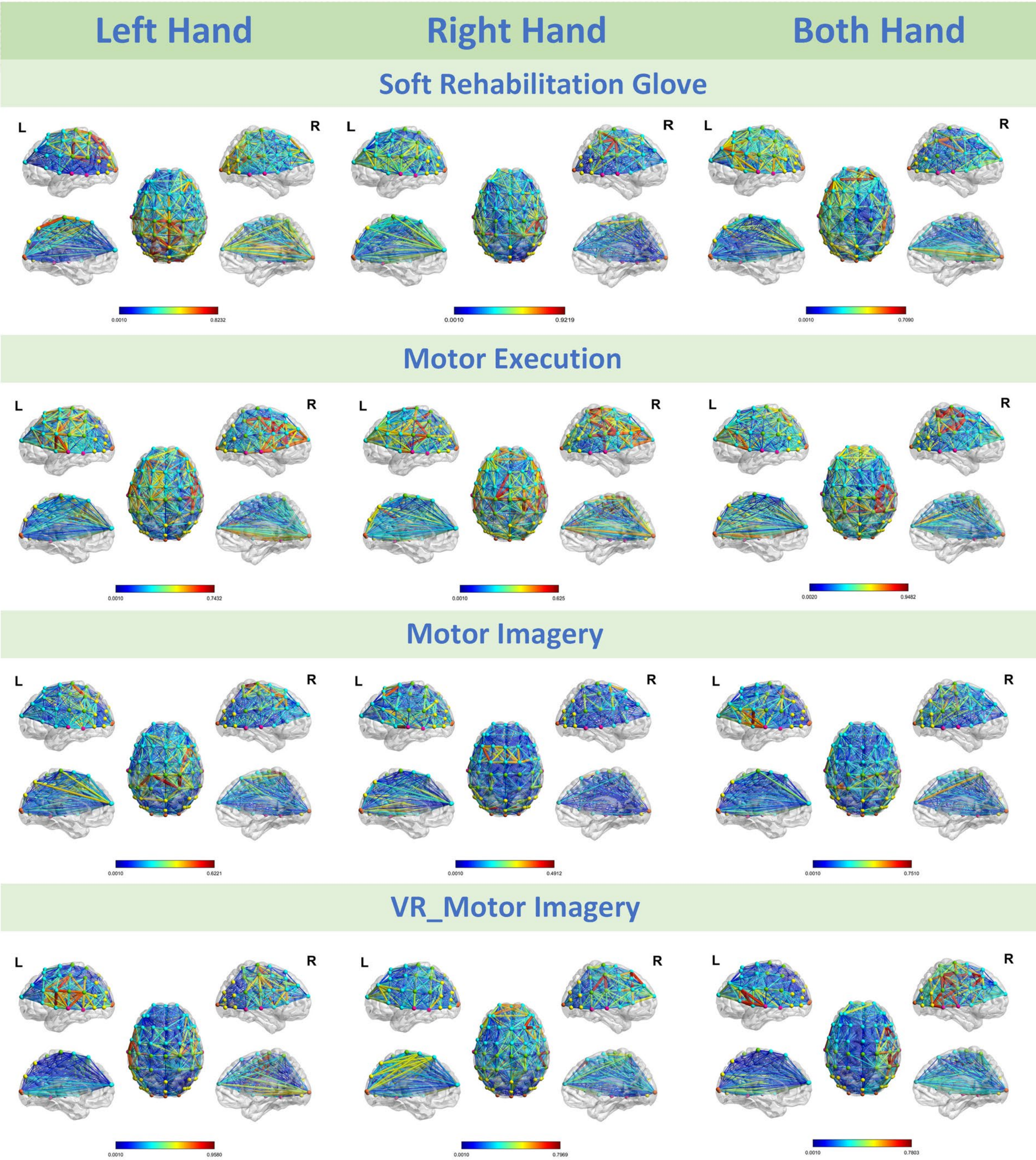

Fig. 15 Brain Functional Network Connectivity Diagrams 2.

Brain functional network diagram. This section presents the brain functional network diagrams corresponding to the execution of six types of experimental paradigms. The calculation method is as follows, for all subjects, we calculated the Phase Lag Index (PLI) for epochs under various paradigms. This index is commonly used to measure the asymmetry in the distribution of phase differences between two signals. It reflects the consistency of phase lead or lag between two signals and serves as a useful measure of phase synchronization.

After obtaining the PLI matrix, we used the 59 electroencephalography (EEG) electrode locations as network nodes and the values representing functional connectivity strength between nodes in the PLI matrix as edge weights. The connection weights in the matrix were transformed into a weight vector to clarify the connectivity relationships between nodes, ultimately completing the construction of the brain functional network.

As shown in Figs. 14, 15, nodes of different colors in the network represent distinct brain regions (blue: frontal lobe; light blue: central region; green: parietal lobe; orange: occipital lobe; red: temporal lobe). Variations in line color and thickness indicate the strength of connections within the brain network. Revealing distinct patterns across the six types of experimental paradigms in terms of connectivity among the primary motor cortex,

	SRG1 ME1	SRG2 ME2	SRG3 ME3	ME1 MI1	ME2 MI2	ME3 MI3
S201	—	**	*	—	—	—
S205	—	—	**	—	—	—
S219	—	—	*	—	—	—
	MI1 VR1	MI2 VR2	MI3 VR3	ME1 VR1	ME2 VR2	ME3 VR3
S201	**	—	*	*	—	*
S205	**	*	*	**	—	—
S219	*	—	**	**	—	**

Table 3. Differences between paradigms for the same task. - means no significant difference, *means $p < 0.05$, **means $p < 0.01$.

	SRG1 SRG2	SRG1 SRG3	SRG2 SRG3	ME1 ME2	ME1 ME3	ME2 ME3	IMA1 IMA2
S201	—	—	—	*	—	—	—
S219	—	—	*	—	—	—	—
S227	—	—	—	*	—	*	*
	MI1 MI2	MI1 MI3	MI2 MI3	VR1 VR2	VR1 VR3	VR2 VR3	MIR1 MIR2
S201	—	—	—	**	—	*	—
S219	—	*	—	**	—	**	**
S227	*	—	—	—	—	—	**

Table 4. Differences between tasks under the same paradigm. - means no significant difference, *means $p < 0.05$, **means $p < 0.01$.

premotor area, supplementary motor area, somatosensory cortex, and visual cortex. Notably, the connectivity during actual motor execution was found to be stronger than that during motor imagery tasks. Furthermore, connectivity during actual motor execution exhibited global coordination and regulation across the entire brain, whereas motor imagery tasks primarily involved regulation of the motor cortex, with weaker connectivity in the somatosensory cortex. However, the results also revealed that assisted movement tasks—specifically those involving soft rehabilitation gloves—enhanced connectivity in the somatosensory cortex.

For further analysis, we binarized the weighted matrix using the upper quartile point of each PLI matrix as the threshold, transforming the connection weights into binary form (values above the threshold were set to 1, otherwise 0), thereby constructing a binarized brain functional network. On this basis, we selected the assortativity coefficient (ASS: referring to the tendency of nodes with similar degree values to connect within a network, which can assess the correlation of node degrees in the network), a commonly used network characteristic parameter in complex network analysis, for calculation. Finally, we performed T-test analysis on this parameter across different paradigms or tasks.

The following two tables represent the analysis of differences in ASS parameters under different paradigms (e.g., motor execution vs. motor imagery) or different tasks (e.g., left hand vs. right hand). (SRG: Soft Rehabilitation Glove, ME: Motor Execution, MI: Motor Imagery, VR: VR_Motor Imagery, IMA: Mirror Glove, MIR: Mirror Therapy, 1:left hand, 2:right hand, 3:both hand, SRG1 ME1 means: Soft Rehabilitation Glove with the left hand vs. Motor Execution with the left hand). The results of the 28 subjects are provided in Supplementary File 1 (Tables 3, 4).

Usage Notes

Users interested in the dataset can register on the platform and download the files locally. When using the dataset, please note the following: Data from subjects S207 and S218 were excluded. So, although the serial numbers are S201 to S230, there are actually only 28 subjects.

Data availability

The two types of datasets (raw EEG data and preprocessed EEG data) can be downloaded from Figshare. The link for the raw data is: https://doi.org/10.6084/m9.figshare.28831730.v2; the link for the preprocessed data is: https://doi.org/10.6084/m9.figshare.28837016.v2.

Code availability

This study provides MATLAB code preprocessing_MI for data preprocessing, which can be run directly for data processing and analysis. The code can be found in the preprocessed EEG dataset.

Received: 16 May 2025; Accepted: 20 October 2025;
Published online: 28 November 2025

References

1. Martin, S. S. *et al.* Heart Disease and Stroke Statistics: A Report of US and Global Data From the American Heart Association. *Circulation* (2025).

2. Feigin, V. L. *et al*. World Stroke Organization (WSO): Global Stroke Fact Sheet 2025. *International journal of stroke: official journal of the International Stroke Society* **20**(2), https://doi.org/10.1177/17474930241308142 (2024).
3. Cervera, M. A. *et al*. Brain-computer interfaces for post-stroke motor rehabilitation: a meta-analysis. *Annals of clinical and translational neurology* **5**(5) 651–663 (2018).
4. Zhang, M. *et al*. Efficacy of brain-computer interfaces on upper extremity motor function rehabilitation after stroke: A systematic review and meta-analysis. *NeuroRehabilitation* **54**(2), 199–212 (2024).
5. Yang, W. *et al*. The Effect of Brain–Computer Interface Training on Rehabilitation of Upper Limb Dysfunction After Stroke: A Meta-Analysis of Randomized Controlled Trials. *Frontiers in Neuroscience* **15**, 766879-766879 (2022).
6. Bin, H. *et al*. Electrophysiological Brain Connectivity: Theory and Implementation. *IEEE transactions on bio-medical engineering* **66**(7), 2115-2137 (2019).
7. Logothetis, N. K. *et al*. Neurophysiological investigation of the basis of the fMRI signal. *Nature* **412**, 150–7 (2001).
8. Tak, S. & Ye, J. C. Statistical analysis of fNIRS data: A comprehensive review. *NeuroImage* **85**, 72–91 (2014).
9. Varbu, K., Muhammad, N. & Muhammad, Y. Past, Present, and Future of EEG-Based BCI Applications. Sensors **22**(9), 3331–3331 (2022).
10. Foong, R. *et al*. Assessment of the Efficacy of EEG-Based MI-BCI With Visual Feedback and EEG Correlates of Mental Fatigue for Upper-Limb Stroke Rehabilitation. *IEEE transactions on bio-medical engineering* **67**(3), 786–795 (2020).
11. Douibi, K. *et al*. Toward EEG-Based BCI Applications for Industry 4.0: Challenges and Possible Applications. *Frontiers in Human Neuroscience* **15**, 705064–705064 (2021).
12. Vourvopoulos, A. *et al*. Efficacy and Brain Imaging Correlates of an Immersive Motor Imagery BCI-Driven VR System for Upper Limb Motor Rehabilitation: A Clinical Case Report. *Frontiers in Human Neuroscience* **13**, 244 (2019).
13. Zhang, W. *et al*. Examining the effectiveness of motor imagery combined with non-invasive brain stimulation for upper limb recovery in stroke patients: a systematic review and meta-analysis of randomized clinical trials. *Journal of NeuroEngineering and Rehabilitation* **21.1**, 209–209 (2024).
14. Rimbert, S. *et al*. Impact of the baseline temporal selection on the ERD/ERS analysis for Motor Imagery-based BCI. *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. vol. 2023, 1–4 (2023).
15. Li, F. *et al*. The Dynamic Brain Networks of Motor Imagery: Time-Varying Causality Analysis of Scalp EEG. *International journal of neural systems* **29**(1), 1850016 (2019).
16. Heinrich, C. *et al*. Feasibility and psychophysical effects of immersive virtual reality-based mirror therapy. *Journal of NeuroEngineering and Rehabilitation* **19**(1), 107–107 (2022).
17. Fernández-Solana, J. *et al*. Efficacy of a Rehabilitation Program Using Mirror Therapy and Cognitive Therapeutic Exercise on Upper Limb Functionality in Patients with Acute Stroke. *Healthcare* **12**(5) (2024).
18. Leong, S. C. *et al*. Examining the effectiveness of virtual, augmented, and mixed reality (VAMR) therapy for upper limb recovery and activities of daily living in stroke patients: a systematic review and meta-analysis. *Journal of NeuroEngineering and Rehabilitation* **19**(1), 93 (2022).
19. Blanco Mora, D. A. *et al*. Impact of age, VR, immersion, and spatial resolution on classifier performance for a MI-based BCI. *Brain-Computer Interfaces* **9**(3), 169–178 (2022).
20. Sui, M. *et al*. A soft-packaged and portable rehabilitation glove capable of closed-loop fine motor skills. *Nature Machine Intelligence* **5**(10), 1149–1160 (2023).
21. Rho, E. *et al*. Multiple Hand Posture Rehabilitation System Using Vision-Based Intention Detection and Soft-Robotic Glove. *in IEEE Transactions on Industrial Informatics* **20**(4), 6499–6509 (2024).
22. Zhang, R. *et al*. Hybrid Brain-Computer Interface Controlled Soft Robotic Glove for Stroke Rehabilitation. *IEEE journal of biomedical and health informatics* (2024).
23. Chen, X. *et al*. A Wearable Hand Rehabilitation System with Soft Gloves. *IEEE Transactions on Industrial Informatics PP.* **99**, 1–1 (2020).
24. Liu, H. *et al*. An EEG motor imagery dataset for brain computer interface in acute stroke patients. *Sci Data* **11**, 131 (2024).
25. Ma, X., Qiu, S. & He, H. Multi-channel EEG recording during motor imagery of different joints from the same limb. *Sci Data* **7**, 191 (2020).
26. Liu, Y. *et al*. Lower limb motor imagery EEG dataset based on the multi-paradigm and longitudinal-training of stroke patients. *Sci Data* **12**, 314 (2025).
27. Wirawan, I. M. A. *et al*. Acquisition and processing of Motor Imagery and Motor Execution Dataset (MIMED) for six movement activities. *Data in Brief* **56**, 110833–110833 (2024).
28. Shuqfa, Z., Lakas, A. & Belkacem, A. N. Increasing accessibility to a large brain–computer interface dataset: Curation of physionet EEG motor movement/imagery dataset for decoding and classification. *Data in Brief* **54**, 110181 (2024).
29. Ma, J. *et al*. A large EEG dataset for studying cross-session variability in motor imagery brain-computer interface. *Sci Data* **9**, 531 (2022).
30. Ganguly, K. *et al*. Modulation of neural co-firing to enhance network transmission and improve motor function after stroke. *Neuron* **110**(15), 2363–2385 (2022).
31. Barry, A. J. *et al*. Survivors of Chronic Stroke Experience Continued Impairment of Dexterity But Not Strength in the Nonparetic Upper Limb. *Archives of physical medicine and rehabilitation* **101**(7), 1170–1175 (2020).
32. Desrosiers, J. *et al*. Performance of the 'unaffected' upper extremity of elderly stroke patients. *Stroke* **27**(9), 1564–70 (1996).
33. Guo, N. *et al*. SSVEP-Based Brain Computer Interface Controlled Soft Robotic Glove for Post-Stroke Hand Function Rehabilitation. *IEEE Trans Neural Syst Rehabil Eng.* **30**, 1737–1744 (2022).
34. Bönstrup, M. *et al*. Low-Frequency Brain Oscillations Track Motor Recovery in Human Stroke. *Annals of neurology* **86**(6), 853–865 (2019).
35. Chang, W. *et al*. A Multi-Paradigm EEG Dataset for Studying Upper Limb Rehabilitation Motor Raw EEG Dataset. *Figshare* https://doi.org/10.6084/m9.figshare.28831730.v2 (2025).
36. Chang, W. *et al*. A Multi-Paradigm EEG Dataset for Studying Upper Limb Rehabilitation Exercises Preprocessed Dataset. *Figshare* https://doi.org/10.6084/m9.figshare.28837016.v2 (2025).
37. Cho, H., Ahn, M., Ahn, S., Kwon, M. & Jun, S. C. EEG datasets for motor imagery brain–computer interface. *GigaScience*, **6**(7) (2017).

Acknowledgements
This research was funded by several sources, including: the National Natural Science Foundation of China under grant 62366028, grant 62466032, and grant W2421090;(ii) Natural Science Foundation of Gansu Province under grant 24JRRA256;(iii) Major science and technology projects of Gansu Province under grant 23ZDFA012;(iv) Open project of Neuracle Company under grant BRKOT-LZJTU-20220329C;(v)The Royal Society (UK) through the International Exchanges 2024 Cost Share Programme (NSFC), under award IEC\NSFC\242277, supporting collaborative work between the University of Essex and Lanzhou Jiaotong University.

Author contributions

W.W.C.: supervision, visualization, methodology, writing-reviewing, resources, funding acquisition. W.X.K.: methodology, collected, analysed the data, writing-editing, and reviewing. G.H.Y.: supervision, resources, funding acquisition. R.J.L., K.Y.D.: collected the data, writing- reviewing. M.T.S., B.G.: Resources, Conceptualization. R.Y., X.L.: Supervision, Resources, Conceptualization.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41597-025-06147-6.

Correspondence and requests for materials should be addressed to W.C.

Reprints and permissions information is available at www.nature.com/reprints.