



OPEN Partial prior transfer learning based on self-attention CNN for EEG decoding in stroke patients

Jun Ma^{1✉}, Wanlu Ma³, Jingjing Zhang¹, Yongcong Li⁴, Banghua Yang^{4✉} & Chunlei Shan^{1,2✉}

The utilization of motor imagery-based brain-computer interfaces (MI-BCI) has been shown to assist stroke patients activate motor regions in the brain. In particular, the brain regions activated by unilateral upper limb multi-task are more extensive, which is more beneficial for rehabilitation, but it also increases the difficulty of decoding. In this paper, self-attention convolutional neural network based partial prior transfer learning (SACNN-PPTL) is proposed to improve the classification performance of patients' MI multi-task. The backbone network of the algorithm is SACNN, which accords with the inherent features of electroencephalogram (EEG) and contains the temporal feature module, the spatial feature module and the feature generalization module. In addition, PPTL is introduced to transfer part of the target domain while preserving the generalization of the base model while improving the specificity of the target domain. In the experiment, five backbone networks and three training modes are selected as comparison algorithms. The experimental results show that SACNN-PPTL had a classification accuracy of $55.4\% \pm 0.17$ in four types of MI tasks for 22 patients, which is significantly higher than comparison algorithms ($P < 0.05$). SACNN-PPTL effectively improves the decoding performance of MI tasks and promotes the development of BCI-based rehabilitation for unilateral upper limb.

Keywords Motor imagery, Transfer learning, Self-attention, Convolutional neural network, EEG decoding

Stroke is associated with a high incidence of disability and mortality^{1,2}. Currently, the effectiveness of rehabilitation treatment is limited for most patients³. Traditional passive rehabilitation methods face challenges in activating the patient's motor nerve center and promoting neural remodeling. The use of motor-imagery based brain-computer interface (MI-BCI)^{4–6} has shown promise in enabling stroke patients to proactively engage their brain's motor regions, thus enhancing outcomes in rehabilitation training^{7,8}. During MI-BCI rehabilitation training, neurons around the damaged area of the patient's brain are repeatedly stimulated, facilitating the remodeling of new neural circuits. While the current motor imagery (MI) tasks involve bilateral limb movements, such as grasping, it is worth noting that patients typically experience unilateral limb paralysis. Consequently, the rehabilitation value of motor imagery involving healthy limbs is limited for these patients. Achieving optimal rehabilitation outcomes becomes challenging when only one type of task is applied to the unilateral limb, as it fails to activate a broad range of motor brain area effectively⁹. Therefore, the precise classification of multiple tasks for the unilateral upper limb holds significant importance for rehabilitation training.

Our past studies^{9,10} introduced four types of unilateral upper limb motor imagery tasks, and both studies confirmed the distinguishability between these tasks. Decoding multiple motor imagery tasks of unilateral upper limb poses challenges, even when utilizing electroencephalogram (EEG) data from healthy subjects^{11–13}. The spatial feature differences in motor brain regions are weakened by the damaged brain in patients, exacerbating the difficulty of decoding the patient's motor intentions¹⁴. Consequently, algorithms for motor imagery face a heightened challenge in decoding the more intricate unilateral upper limb motor imagery tasks.

Reviewing the recent advancements in motor imagery decoding, there has been a transition from traditional machine learning to deep learning, marked by an expansion in the number of tasks and an improvement in classification accuracy. Classic traditional machine learning algorithms, such as feature extraction methods like common spatial patterns (CSP)¹⁵ and filter bank common spatial patterns (FBCSP)¹⁶, have evolved gradually to achieve more precise classification for both two-class and multi-class tasks. Traditional machine learning, with its

¹Department of Rehabilitation Medicine, Tong Ren Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 200336, China. ²Institute of Rehabilitation, Shanghai Jiao Tong University School of Medicine, Shanghai 200025, China. ³China-Japan Friendship Hospital, Beijing 100029, China. ⁴School of Mechatronic Engineering and Automation, Shanghai University, Shanghai 200444, China. ✉email: mj1990@shu.edu.cn; yangbanghua@shu.edu.cn; shancl@shsmu.edu.cn

lower computational overhead and shorter modeling time, aligns more closely with the modeling requirements of online BCI, making it widely applicable in BCI rehabilitation and control systems^{17,18}. In recent years, to address the challenges of algorithm classification accuracy and complexity, hybrid deep learning has gained significant attention due to its superior spatiotemporal mining capabilities and shorter hardware processing times¹⁹. Notable research in this area includes the integration of convolutional neural networks with recurrent neural networks²⁰ and with deep belief networks²¹. However, with the advancement of computing power, deep learning algorithms have gradually been applied to online BCI systems, evolving alongside advanced network architectures. Among the most effective deep learning network structures are EEGNet²², FBCNet²³, and Deep ConvNets²⁴. With the development of deep learning, state-of-the-art EEG decoding algorithms have been proposed such as ATCNet²⁵, EEG conformer²⁶, GCNs-Net²⁷, Tensor-CSPNet²⁸, and BrainGridNet²⁹ which have more advanced network structures and higher decoding performance. However, it is worth noting that the classification accuracy is typically enhanced by only 3–5% points, primarily due to the limited generalizability resulting from the small sample size of EEG data when applying advanced models.

The advent of transfer learning has opened up new possibilities for enhancing the generalizability of deep network models. In this approach, EEG data from other subjects (within the same dataset as the target domain) serves as the source domain for transfer learning, leading to improved performance by identifying generalized features between the source domain and the target subject data (target domain). Throughout the training process, emphasis is placed on ensuring generalization across each class of tasks, learning the specific features of the target subjects, and enhancing the adaptability and discriminative ability of each class of specific features³⁰. Presently, significant performance gains are observed in the EEG decoding problem through transfer learning^{31–35}, even though the structures of these networks may not be highly complex. This suggests that the collaborative application of transfer learning and deep learning is likely to further enhance EEG decoding abilities.

This paper introduces a self-attention convolutional neural network (SACNN) tailored to the spatial decoding properties of EEG. SACNN enhances the correlation between different channels (brain region locations) and elevates the weights of channels relevant to the task. SACNN treats each EEG channel as a sentence and leverages the self-attention mechanism to calculate the correlations between these sentences, the correlations between channels. This enables the EEG channels to have their weights reallocated. During this process, channels with stronger correlations are assigned higher weights, which means that brain regions with increased activity during task execution receive greater emphasis. These regions contribute most significantly to task decoding, as they exhibit changes directly influenced by the task. Additionally, partial prior transfer learning (PPTL) is proposed to augment the model's generalization ability. The partial prior, representing a portion of the training set from the target domain, is integrated into the training of the source domain model as a validation set. This integration ensures that the model output aligns more closely with the features of the target domain, striking a balance between maintaining generalization from the source domain and specificity to the target domain. The paper concludes with a discussion on the influence of the amount of partial prior on the model and the performance of self-attention convolutional neural network based partial prior transfer learning (SACNN-PPTL).

Materials and methods

Dataset

Stroke patients typically experience unilateral limb paralysis, particularly affecting the hand and upper limb. The brain regions governing the motor functions of these areas are extensive. Neuroplasticity is potentially enhanced when these brain regions are activated as much as possible. Therefore, four types of MI tasks for unilateral upper limb are designed in this paper to activate the motor brain regions of patients in a wider range. Based on previous studies on healthy subjects^{9,10}, MI data from stroke patients is used as research objects to study the classification algorithm and the deployment scheme. The dataset collected EEG data for four types of MI from 22 stroke patients. Previous research examined the classification accuracy for some subjects within this dataset³⁶, demonstrating the distinguishability of the four types of tasks. However, there is minimal difference in brain activation distribution between the different tasks. The brain map of α and β frequency bands during the four types of motor imagery tasks for all subjects are shown in Fig. 1. The brain map of all subjects exhibit similar event-related desynchronization/synchronization (ERD/S) characteristics, with smaller feature differences and greater decoding difficulty compared to the bilateral MI tasks for the left and right hands.

Experimental paradigm

The MI tasks and experimental paradigm are depicted in Fig. 2. The four categories of MI tasks in Fig. 2(a) include grasping (class 1), catching (class 2), reaching (class 3), and rotation (class 4). For grasping, subjects are instructed to open and close their hand; catching requires subjects to move an object horizontally for a certain distance and then return it; reaching involves extending the hand forward and then retracting it; rotation requires the wrist to rotate clockwise by 180° and then return counterclockwise.

Before the experiment, subjects wore the EEG headset. The state of the electrodes was adjusted until the impedance dropped below the threshold value. At the beginning of the experiment, subjects were instructed to sit comfortably in a chair positioned 1 m from the screen. The experimental paradigm is illustrated in Fig. 2(b). Initially, the screen displayed a cue for the resting state, with eyes open for one minute followed by one minute of eyes closed. Subsequently, the screen randomly presented an MI task with a 3-second voice prompt. Subjects were required to repeatedly imagine the task during the task period (4-second) while refraining from any physical movements. The performance of the MI task was sustained until the screen turned black, accompanied by an audible “rest” signal. Following a 2-second rest, the trial concluded. Each subject participated in five sessions, with each session consisting of 40 trials. These five sessions were conducted consecutively in the same experiment. At the end of each session, subjects were allowed to rest for one minute or more.

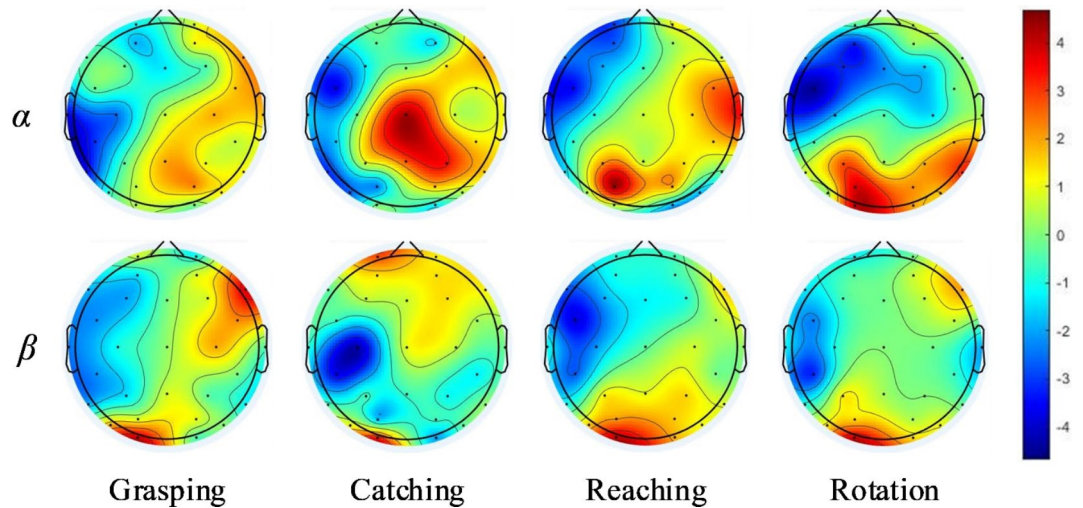


Fig. 1. Brain maps of each task in α and β band, where red indicates an increase in EEG energy and blue indicates a decrease in EEG energy.

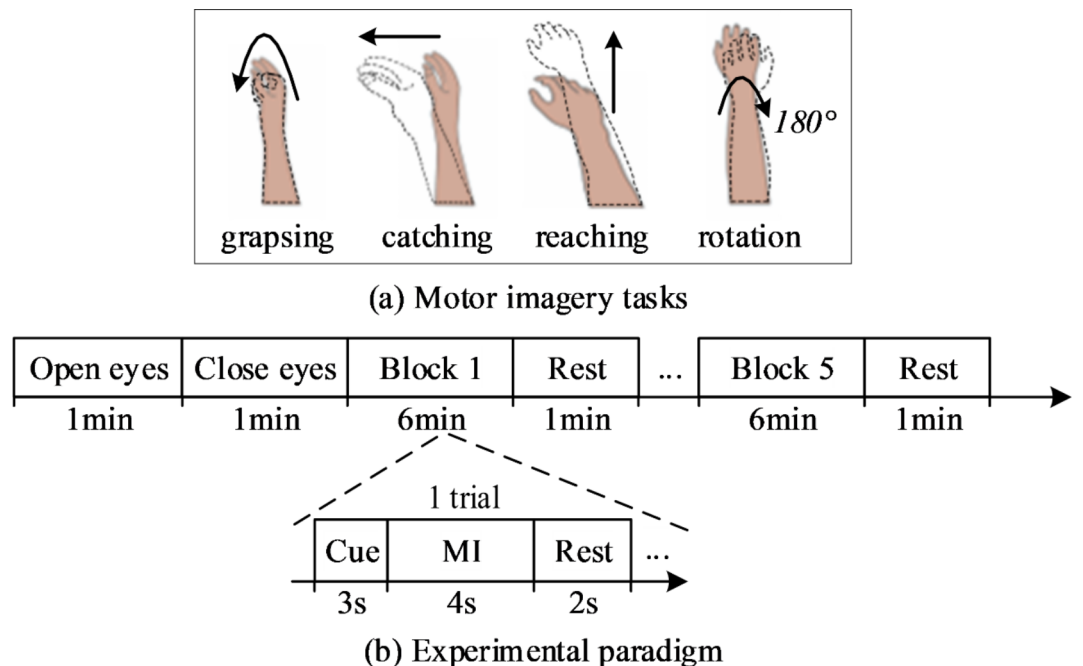


Fig. 2. MI tasks and experimental paradigm. (a) MI tasks: Four movements represent four types of MI tasks for unilateral upper limb. (b) Experimental paradigm: The complete timeline of one experiment, where a trial extended from Block 1 represents the smallest unit making up a block, with each block consisting of several trials.

Data acquisition and pre-processing

22 stroke patients (5 females, age 65.9 ± 9.1 years) participated in this study, all of whom are right-handed. All subjects had left upper extremity paralysis with a Brunnstrom score of 1–5³⁷. Before the experiment, all subjects are familiar with the experimental content and all precautions. The study has been approved by the Human Research Ethics Committee of Shanghai Gongli Hospital (GLYYIs2021-025) and is in accordance with the Declaration of Helsinki. This study received informed consent from all participants and/or their legal guardians, with all participants signing the informed consent form (ICF) prior to the start of the research.

The EEG headset (CGX, American) is equipped with 32 dry electrodes designed for data acquisition at a sampling rate of 1000 Hz. This equipment features amplifier-integrated electrode caps, acquisition, and recording software. The analog-to-digital conversion is carried out with a 24-bit simultaneous sampling, and the noise level is $0.7 \mu\text{V}$ Root Mean Square (RMS) within the frequency range of 1–50 Hz.

Data acquisition involves external interference signals, such as eye movement, electrocardiogram, and electromyographic artifacts³⁸. The interference components in EEGLAB (<http://sccn.ucsd.edu/eeGLAB>) are removed through independent component analysis (ICA). After ICA analysis, the ADJUST toolbox is used to automatically label and remove noise components (ocular artifacts, muscle artifacts, etc.)³⁹. Trials with large interference which can be observed manually are removed. The EEG data is down sampled to 250 Hz to enhance the signal-to-noise ratio while reducing computational load, as the majority of the electrophysiological signals (below 100 Hz) can be fully retained^{23,24,40}.

Self-attention convolutional neural network based partial prior transfer learning

Model-based transfer learning demonstrates superior classification performance in MI decoding because the pre-trained model (base model), which includes a large number of subject-specific EEG features, possesses high generalizability³⁰. After training the base model with a large amount of source domain data, the model learns broader features across different tasks, which are then transferred to the target domain through transfer learning. As a result, the model exhibits greater generalization capability compared to the within mode, which only uses target subject. Training base model is a common and effective method to reduce the training sets of target domain in computer vision modeling⁴¹. Nevertheless, transfer learning is still difficult to take full advantage of in EEG decoding, because EEG is very specific, even if the within-subject shows large differences in characteristics across different brain states, let alone between-subject. Therefore, features that match the specificity of the target domain are mined in the source domain to help the basic model retain the generalization of large samples and improve the specificity of the target domain.

The SACNN-PPTL for priorly testing the specificity of the target domain is proposed. The framework includes a source domain composed of data from many patients and a target domain composed of data from a single patient, as shown in Fig. 3. In the source domain these percentages represent the proportion of training samples, the source domain data is divided into 90% training sets and 10% validation sets. In the target domain, all samples of the target subject are divided into the training set, verification set, and the test set according to the ratio of 80%, 10%, and 10%. Since the target domain is trained by one subject each time, 10-fold cross-validation is carried out for each subject according to the sample size ratio of 8:1:1. Partial prior is first introduced to train the base model, then the base model is transferred to the target domain for retraining, and finally the retrained model is used to predict the test set. The base model in Fig. 3 is trained by the source domain represented by blue and the partial prior represented by orange in the target domain, and its basic network is SACNN. The network parameters for each layer of the base model, the loss calculation, and the optimizer backpropagation are shown in blue. In the target domain, except for the temporal feature layer, which retains the parameters of the basic model (blue), the other network layers are updated according to the target domain training, and the final target model is tested by the test set of target domain.

Suppose the input of SACNN-PPTL is X^i , s , p , and t representing the source domain, partial prior, and target domain, respectively. $X_t^i \in R^{N_c \times N_t}$ represents one trial of the target domain, N_c represents the number of EEG channels, and N_t represents time samples. The model of transfer learning can be regarded as a classifier: $f: R^{N_c \times N_t} \rightarrow L$, $R^{N_c \times N_t}$ represents an EEG trial, N_c represents the number of EEG channels, and N_t represents time samples, L represents the four types of tasks (labels). Suppose f_t is the base model trained by the target domain. The mathematical model of SACNN-PPTL can be expressed as Eq. (1).

$$f_t(X_t^i, \theta') = g'(\phi'(X_t^i; \theta_{f_1}, \theta'_{\phi_2}, \dots, \theta'_{\phi_N}); \theta'_g) \quad (1)$$

where θ_{f_1} denotes the network parameters of the first convolutional layer of the fixed source domain model, θ'_{ϕ_i} denotes the training parameters of the other convolutional layers, ϕ' is the network structure trained in the target domain, and g' denotes the softmax in the target domain.

Spatial features are not only the main basis for decoding, but also the difference of task features among different subjects. The space-sensitive network will facilitate the adaptation of the base model to the target domain in the PPTL. In bilateral limb tasks, the greatest disparity between left-hand and right-hand MI tasks lies in the activation differences in spatial features. Similarly, in multi-task for unilateral upper limb, spatial features are more critical as they provide the most compelling evidence for distinguishing between different tasks in brain activation. Decoding becomes more challenging when event-related desynchronization/synchronization (ERD/S) arises from the multi-task of unilateral upper limb¹³. Traditional convolutional neural networks and deep learning networks with filter band characteristics are insufficient to extract precise spatial feature differences²³. SACNN based on channel self-attention (SA) exhibits unique advantages in decoding such problems⁴². It calculates the correlation weights between each channel after the temporal feature extraction module and enhances the weights of channels with higher correlations. This process amplifies the phenomenon of increased activation on the same side or decreased activation on the opposite side during specific MI tasks at the channel SA layer⁴². Consequently, this enhances the model's learning ability regarding EEG spatial features and improves the model's spatial recognition capability.

Self-attention convolutional neural network

The structure of SACNN contains a temporal feature extraction module (Module 1), a spatial feature extraction module (Module 2), and a feature generalization module (Module 3), as shown by SACNN in Fig. 3. The temporal features of the raw EEG are randomized by some convolutional kernels in the temporal feature extraction module. The channel weights of the EEG are altered in the spatial feature extraction module and channels with higher correlation are assigned larger values. Finally, the temporal and spatial features of EEG are gradually generalized in the feature generalization module and output generalized features.

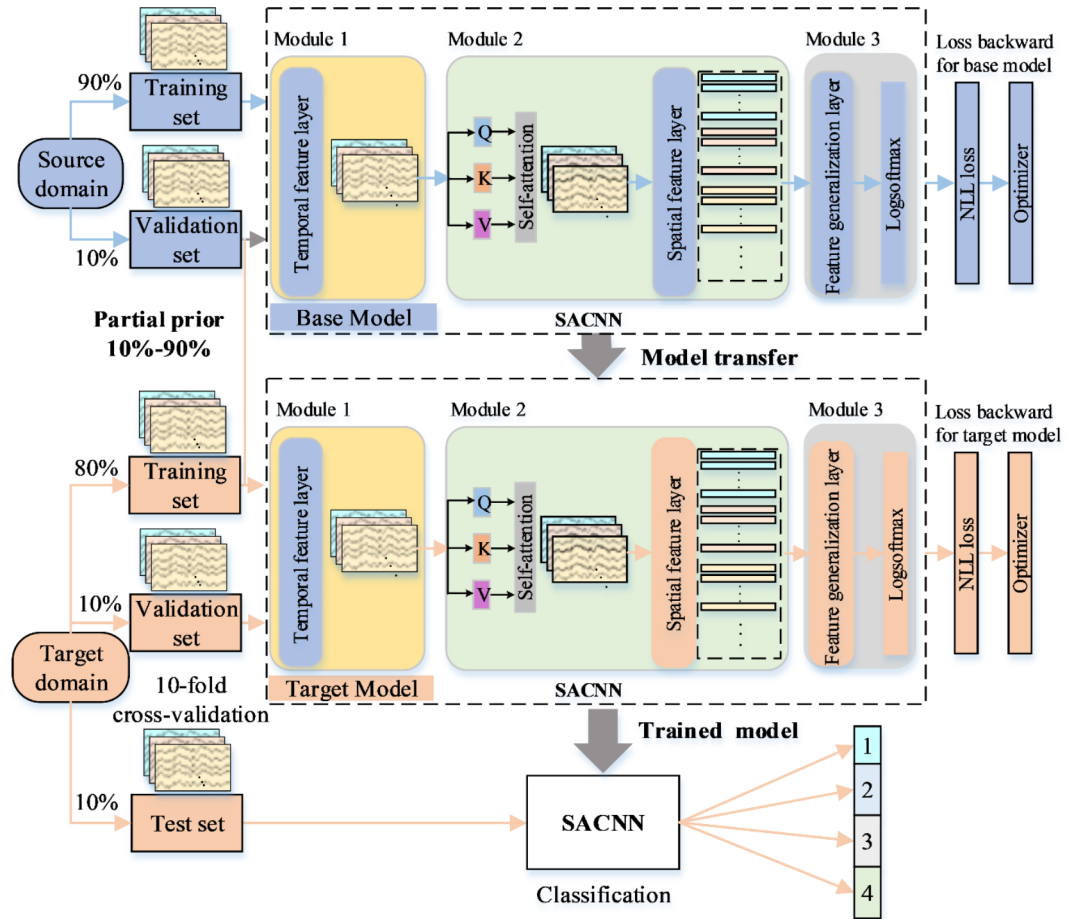


Fig. 3. SACNN-PPTL network structure. The blue part represents the source domain, orange represents the target domain, and arrows indicate the data flow. SACNN is the base network for transfer learning. The base model is trained in the source domain, then the model is transferred to the target domain for target domain adaptive training. The trained model is used for classification on the test set, yielding the test results. In the source domain these percentages represent the proportion of training samples, the source domain data is divided into 90% training sets and 10% validation sets. In the target domain, all samples of the target subject are divided into the training set, verification set, and the test set according to the ratio of 80%, 10%, and 10%. 10-fold cross validation is performed in the target domain.

The temporal feature extraction module is a two-dimensional (2D) convolutional layer⁴³ containing a total of 25²⁴ kernels that stochastically optimize and extract temporal features in the temporal dimension of the EEG. Since the most relevant frequency bands for MI are α and β , the signals in the frequency range 8–30 Hz contain the most effective feature information⁴⁴. EEG signal features below 31.25 Hz can be kept at 250 Hz sampling rate when the kernel size is 1×4 . The convolution kernel size is set to 1×4 at a sampling rate of 250 Hz, allowing the retention of EEG signal components below 31.25 Hz. This is because a 1×4 kernel size can capture the signals within the frequency range covered by four consecutive time points. During the sliding window calculation in convolution, all EEG signal components below 31.25 Hz are captured without omission. The parameter of temporal feature extraction module is as follows. The twenty-five 2D feature maps output by convolutional computation are connected in the depth dimension to form the 3D feature matrix²⁴. Each eigenvalue in the feature matrix is computed by the convolution of Eq. (2).

$$\hat{X}(i, j) = \sum_{m=1}^m \sum_{j=1}^n (X(m, n) \times K_e(i - m, j - n)) \tag{2}$$

where $X(m, n)$ is the input of 2D convolutional, K_e is the kernel size, and $\hat{X}(i, j)$ is the eigenvalue of the output of the convolutional computation, which is at a coordinate position (i, j) in the feature map.

The structure of spatial feature extraction module includes a channel SA layer and a spatial feature layer which include two 2D convolutional layers. The input of the channel SA layer is the feature matrix with depth 25 (output of Module 1). W^Q , W^K , and W^V in the channel SA layer denote three linear layers whose function is to randomly weight the EEG channels without changing the dimension of the time width, so the input and

output dimensions of the linear layers are the same as the time length of the feature matrix. For channel SA computation, the first step is to calculate a score for each vector (EEG channel). The score is obtained by taking the product of the post-randomization X_i matrix queries (Q) and the transpose of keys (K). The normalization of the scores by $\sqrt{d_k}$ is performed to ensure gradient stability. Subsequently, softmax activation is applied to calculate relevance weights based on the scores. Finally, the weights are dot-multiplied with the randomized X_i matrix values (V) to obtain the weighted matrix.

The working principle of channel SA is illustrated in Fig. 4. The method for calculating correlations between sentences in natural language processing is applied to EEG decoding⁴⁵, where each EEG channel is treated as a sentence and each time sample as a word. Therefore, X^i can be decomposed into $\{c_1, c_2, \dots, c_i\}$, where each c_i represents a sentence. The relevance weights between sentences, which represent channel weighting, are obtained by calculating SA between sentences. Each trial of EEG is regarded as a passage in natural language processing⁴⁵, where time samples are sequences of words and channels representing the number of sentences. The correlations between words are computed equivalently to the correlations between EEG channels as shown in Eq. (3).

$$Z^i = softmax \left(\frac{X^i \times W^Q \times (X^i \times W^K)^T}{\sqrt{d_k}} \right) X^i \times W^V \tag{3}$$

where i represents the EEG channel, W represents the weight matrix, X^i denotes a trial of EEG, d_k is the dimension of matrix K, $\sqrt{d_k}$ is used for normalization computation, and the output Z^i is the EEG weighted by SA. The channels of the feature matrix are weighted through the channel SA layer, enhancing the correlation between the EEG channels and the spatial specificity of the output features.

The first 2D convolutional layer is used for weight calculation of spatial features, and the input is a feature matrix of 25 depth, which is randomly optimized and computed in the channel dimension using 25 kernels for 2D convolution. The kernel size is 1×4 and the stride is 1. The size of the convolution kernel is determined by the relationship between the EEG channels. When the channel convolution kernel is 4, at least two neighboring channels are computed at a time, depending on the distribution of EEG electrodes and the method of data storage. Adjacent channels are combined to compute convolutions based on the similarity of the weights of adjacent EEG channels, which maximizes the weights of highly activated regions of the brain and facilitates network learning in the most active brain regions of the subject. The channel weighted spatial features are extracted in the second 2D convolutional layer. Behind the second convolutional layer are the batch normalization (BN)⁴⁶, exponential linear unit (ELU)⁴⁷, and maximum pooling (MaxPool)⁴³ layers.

The feature generalization module is a feature generalization layer consists of three convolutional pooling layers. Each convolutional pooling layer includes a 2D convolutional layer, and a logarithmic softmax (Logsoftmax) layer. The feature generalization module is used to increase the depth of the network and generalize the features. The features are gradually generalized as the number of kernels increases sequentially on different convolutional layers and the highly generalized feature vector is finally output. The 2D convolutional layer at the

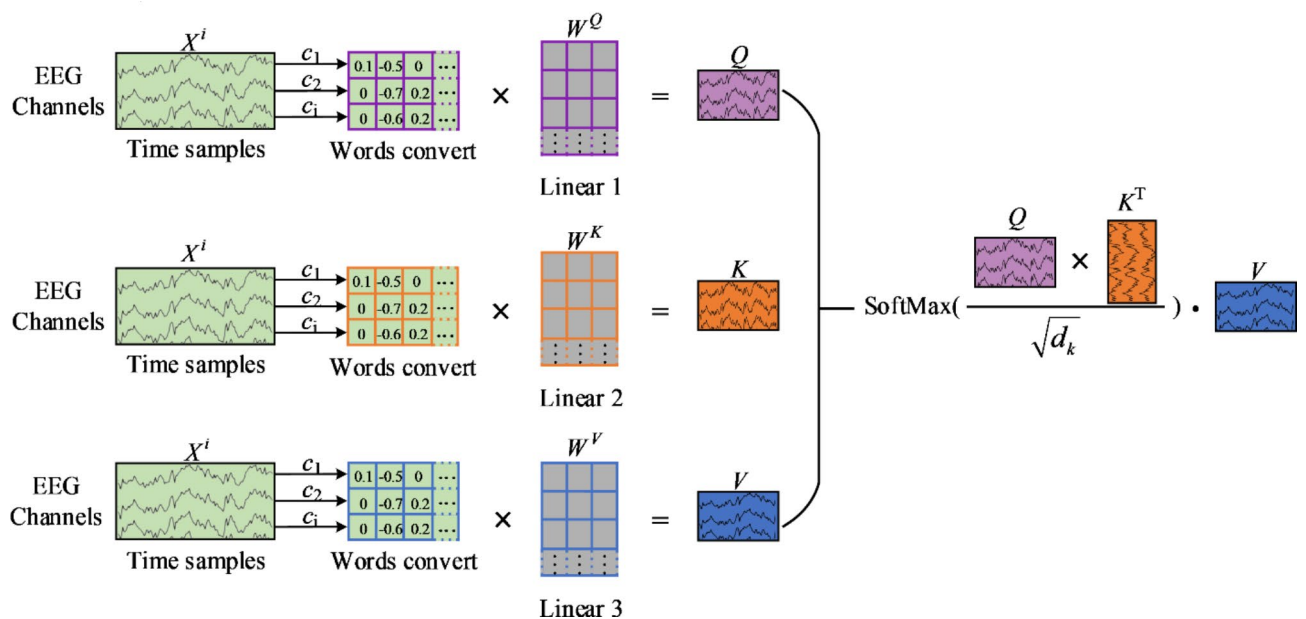


Fig. 4. The computational framework for self-attention. EEG channels are quantified into forms similar to words, with each channel (c_i) representing a sentence. Then, channel-weighted EEG is obtained through SA computation.

end of the module outputs a feature vector of length 4 through 4 kernels. Finally, the features are normalized by Logsoftmax and output. The Dropout layer prevents overfitting during model training. The parameters of the convolutional pooling layer are: kernel size is 1×10 and stride is 1. The number of kernels in the three convolutional layers is 50, 100, and 200 respectively. The pooling kernel size is uniformly set to 1×3 and the stride is 3. The dimensions of the output feature matrix for each layer of the SACNN are shown in Table 1 when the input size of the network is 29×1000 .

Partial prior transfer learning

Model-based transfer learning relies on training a base model using a large-sample dataset. However, task-specific generalization features cannot be effectively learned in EEG decoding of stroke patients. This is because different spatial features come from the tasks of patients with different brain injury regions. Therefore, not all samples in the large dataset exhibit consistent spatial features, leading to the base model's generalization exceeding the data features of the target samples. Training the base model using the target domain can enhance the base model's specificity in the target domain, directing it towards optimizing for the distribution of target domain features. PPTL leverages this transfer concept by using a portion of the target domain training set to constrain the training of the base model. This approach ensures the generalization of the base model while enhancing its specificity to the target domain, thereby improving the final model's classification performance.

The structure of PPTL shown in Fig. 3 refers to adaptive transfer learning³⁰. The partial prior data is used for training on the validation set to improve the performance of the source domain model. The source domain is split into a training set and a validation set. The target domain is split into a training set and a test set, where part of the training set is used as a partial prior. The partial prior and the source domain validation set are recombined to form a new validation set with the source domain training set to train the base model of SACNN. Then, the base model is migrated to the target domain, where the network parameters of the first convolutional layer are fixed, and the rest of the parameters continue to be trained by the training set. Finally, the classification performance of the target model is examined by the test set.

The flow of SACNN-PPTL is shown in Algorithm 1 which includes two parts: base model training and target model training. The new validation set (X^e, y^e) consists of a validation set (X^{Se}, y^{Se}) in the source domain and a partial prior (X^p, y^p) in the target domain. SACNN is trained in the source domain with training set (X^{St}, y^{St}) and validation set (X^e, y^e) . The input samples go through Module1 to Module3 to extract the feature vectors F_f and the negative log-likelihood (NLL) is used to compute the loss of the model. Finally, the model parameters are updated during the backward propagation. The target model is trained using the target domain (X^{Tt}, y^{Tt}) training set, and only the parameters of Module 1 are fixed during the training process, while the rest of the parameters are continuously updated with the backward propagation. The final target model is obtained, after 200 epochs of iterations.

Partial prior data partitioning

In SACNN-PPTL, the proportion of prior knowledge determines the generalization ability of the source domain model and the degree of specificity to the target domain. During the training of the base model, an appropriate proportion of prior knowledge will facilitate optimization towards the feature distribution of the target domain. However, an excessive proportion of prior knowledge restricts the generalization ability of the base model. The optimal proportion of prior knowledge can intuitively inform readers of how much prior knowledge to use when applying this method, eliminating the need for extensive parameter tuning by the reader. Therefore, the construction process of the base model requires optimization of the proportion of prior knowledge, aiming to find a relatively excellent ratio.

Partial prior data partitioning is shown in Fig. 5, where the gray bar indicates the source domain and the yellow bar indicates the target domain. The target data is split into 10 folds in cross-validation, where the green

Module	Layer	Output	Module	Layer	Output
1	Conv1 (1×4)	25,29,997	3	Dropout	50,1,162
	BN	25,29,997		Conv5 (1×10)	100,1,153
	ELU	25,29,997		BN	100,1,153
	MaxPool (1,2)	25,29,498		ELU	100,1,153
2	Linear	25,29,498	MaxPool (1×3)	100,1,51	
	Linear	25,29,498	Dropout	100,1,51	
	Linear	25,29,498	Conv6 (1×10)	200,1,42	
	Conv2 (1×4)	25,29,495	BN	200,1,42	
	Conv3 (29×1)	25,1,495	ELU	200,1,42	
	Dropout	25,1,495	MaxPool (1×3)	200,1,14	
	Conv4 (1×10)	50,1,486	Dropout	200,1,14	
3	BN	50,1,486	Conv7 (1×14)	4,1,1	
	ELU	50,1,486	Logsoftmax	4,1,1	
	MaxPool (1×3)	50,1,162			

Table 1. The network parameters of SACNN.

bar indicates the test set for each fold and the remaining gray bar is the training set. The training set of the target domain is divided into 10 parts to study different proportions of the partial prior data. The partial prior is taken 10%, 30%, 50%, 70%, or 90% each time and is used with the validation set of the source domain (blue bars) to form a new validation set for base model training.

The proposed algorithm of SACNN-PPTL.

Input: EEG samples and true labels of the source domain (X^s, y^s) and target domain (X^t, y^t).

Output: Classification model for the target domain.

Base model training:

1: Splitting the training (X^{Sr}, y^{Sr}) and validation (X^{Se}, y^{Se}) sets of the source domain.

2: Splitting the training (X^{Tr}, y^{Tr}) and test sets of the target domain.

3: $(X^p, y^p) = (X^{Tr}, y^{Tr})[1:\lambda]$

4: $(X^e, y^e) = \{(X^{Se}, y^{Se}), (X^p, y^p)\}$

5: **for** $epoch = 1:200$ **do**:

6: **for** (X_i^t, y_i^t) in $(X^{St}, y^{St}), (X^e, y^e)$ **do**:

7: $F_t = Module\ 1(X_i^t, y_i^t)$

8: $F_s = Module\ 2(F_t)$

9: $F_f = Module\ 3(F_s)$

10: $Loss = L_{NLL}(F_f)$

11: Update Optimizer

12: **end for**

13: **end for**

Target model training:

14: Model transfer

15: Fixed Module1

16: **for** $epoch = 1:200$ **do**:

17: **for** (X_i^t, y_i^t) in (X^{Tr}, y^{Tr}) **do**:

18: $F_t = Module\ 1(X_i^t, y_i^t)$

19: $F_s = Module\ 2(F_t)$

20: $F_f = Module\ 3(F_s)$

21: $Loss = L_{NLL}(F_f)$

22: Update Optimizer

23: **end for**

24: **end for**

25: **Return:** Classification model for the target domain.

Algorithm 1. SDIF. Note: In the source domain, the data is divided into 90% training set and 10% validation set. In the target domain, all samples of the target subject are divided into the training set (80%), verification set (10%) and test set (10%) according to 10-fold cross validation. λ represents the partial prior rate.

Training

The dataset utilized for training the Model encompasses both source and target domains. The source domain is the EEG data of all subjects but the target one. The target domain is the EEG data for the test subject. The training modes are categorized into three: within, adapt, and PPTL. Within uses only the target domain subjects to divide the training set, validation set, and test set independently according to each individual. 90% of the data in the source domain is selected as the training set, while 10% of the data in the source domain is selected as the validation set. Adapt is adaptive transfer learning, which uses only source domain data to train a base model and then transfers to the target domain for model adaptive training. PPTL is the proposed algorithm in this paper. The two transfer learning algorithms have a learning rate of 0.01 for the base model and 0.005 for the target model, the number of epochs for all models is 200, and the batch size is 32. This experiment is run on PyTorch 1.7.1 deep learning environment based on Windows 10 system with Intel Core i9-10900×(3.7 GHz)

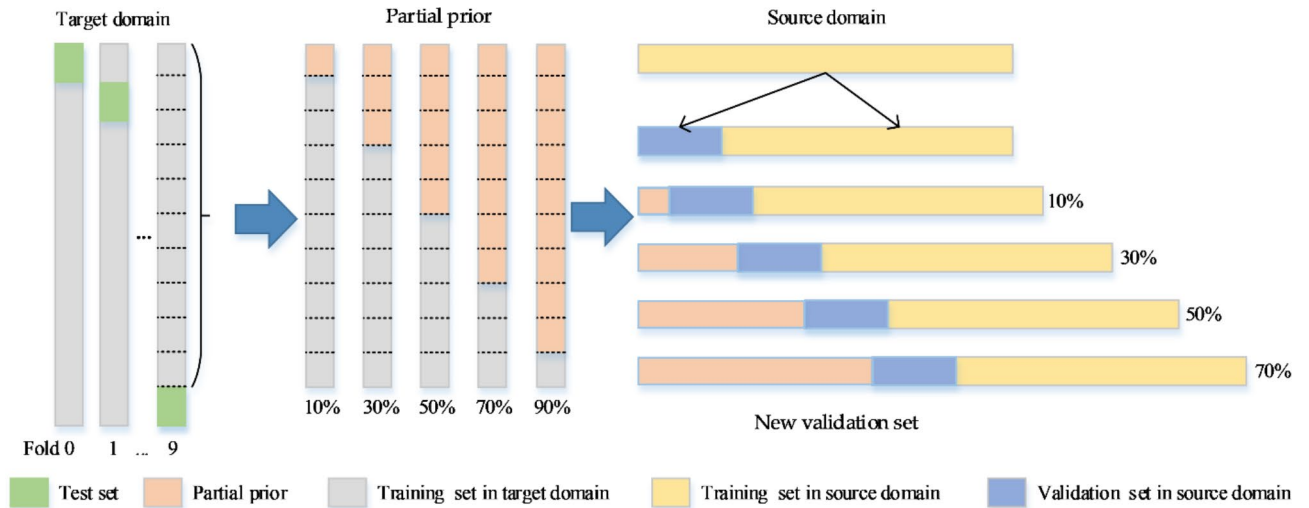


Fig. 5. The partial prior data division is illustrated, where the yellow bars represent EEG data from the source domain, and the gray bars indicate EEG data from the target domain. The target domain is divided into 10-fold cross-validation, with the green bars used as the test set and the gray bars as the validation set.

CPU, 128GB RAM, and NVIDIA GeForce RTX 3090 GPU. The code for this paper is available on GitHub (<https://github.com/JunMa-MIBCI/Partial-Prior-Transfer-Learning-PPTL-/tree/main>).

Evaluation and statistical analysis

Accuracy is a commonly used classification evaluation metric. In both binary and multi-class classification model evaluation, accuracy is a direct evaluation criterion. Apart from accuracy, precision in multi-task classification measures the accuracy of the model in predicting positive samples, while recall measures the model's ability to retrieve positive samples, i.e., the probability of correctly identifying positive samples. F1-score balances the precision and recall of a classification model, providing a more objective reflection of the model's classification performance. In the evaluation of multi-task models, classification accuracy, precision, recall, and F1-score are typically used. The calculation methods for these evaluation metrics are as follows:

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN} \quad (4)$$

$$Precision = \frac{TP}{TP + FP} \quad (5)$$

$$Recall = \frac{TP}{TP + FN} \quad (6)$$

$$F1 \text{ score} = 2 \times \frac{Recall \times Precision}{Recall + Precision} \quad (7)$$

TP represents true positives when both actual and predicted values are true. TN represents true negatives when both actual and predicted values are false. FP stands for false positives, occurring when the actual value is false but predicted as true. FN stands for false negatives, happening when the actual value is true but predicted as false.

In addition, analysis of variance (ANOVA) is employed to compare the differences in classification accuracy between the algorithms proposed in this paper and the benchmark algorithms. Before conducting ANOVA statistical analysis, Kolmogorov-Smirnov and Shapiro-Wilk tests for normality are performed on the accuracy distributions, grouped by algorithm and training mode. After outliers are removed, all accuracy distributions met the normality requirement within a 95% confidence interval. ANOVA is then applied to pairs of results for significance testing, with accuracy as the statistical factor. A p value (P) of less than 0.05 is considered indicative of a statistically significant difference between the classification accuracies of two groups. In this paper, the chance level is determined to be 30% based on the data sample size and total number of classes when significance is set at $P < 0.05$ ⁴⁸. For the comparison of SACNN's classification performance, Deep ConvNets were used as the benchmark algorithm, and ANOVA calculations were conducted across the within, adapt, and PPTL modes. To assess PPTL's classification performance, ANOVA was performed using within and adapt as benchmark algorithms, with SACNN as the base model. For the significance comparison of varying prior accuracy proportions, adapt with SACNN as the benchmark algorithm was compared with PPTL partial priors rate ranging from 10 to 90%.

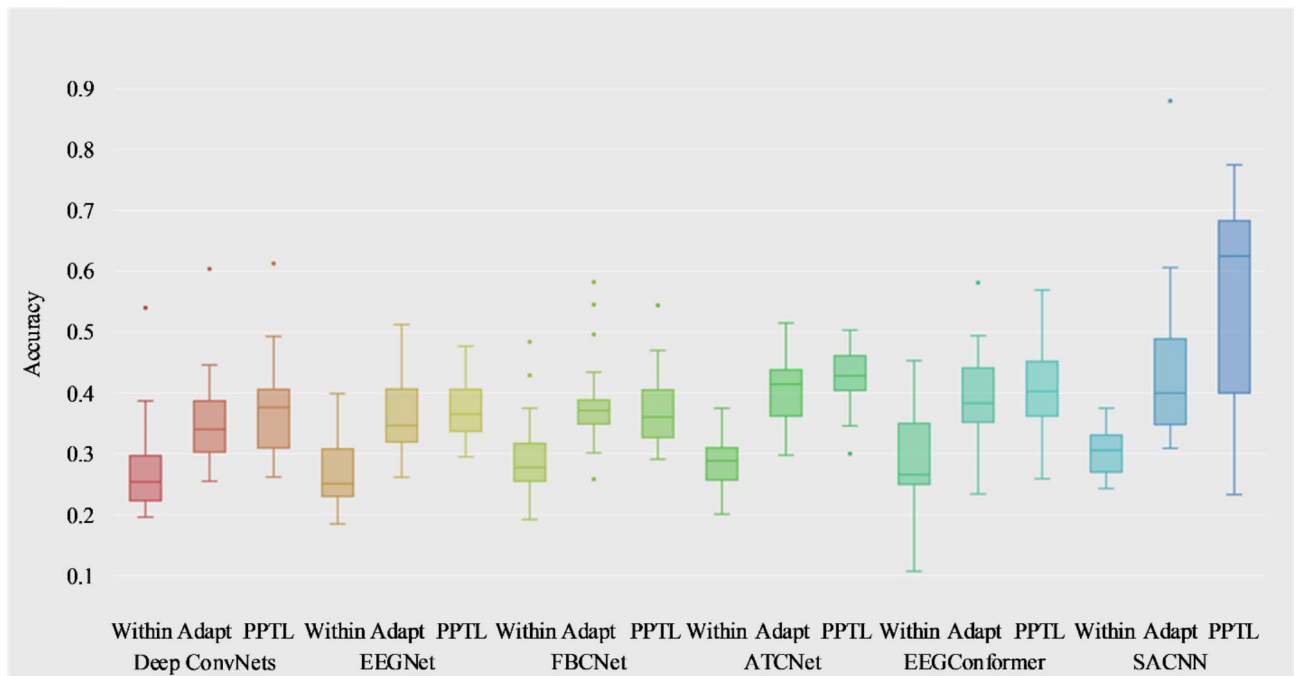


Fig. 6. The accuracy of various networks for each training mode. The partial prior rate is 70% in PPTL. The horizontal axis represents the distribution of classification accuracy for all subjects across different networks and training modes, while the vertical axis indicates the classification accuracy rate.

Algorithm	Deep ConvNets	EEGNet	FBCNet	ATCNet	EEG Conformer	SACNN
Within	0.258 ± 0.08	0.261 ± 0.05	0.286 ± 0.07	0.285 ± 0.05	0.284 ± 0.08	0.305 ± 0.03
Adapt	0.354 ± 0.07	0.362 ± 0.06	0.385 ± 0.07	0.403 ± 0.06	0.397 ± 0.07	0.435 ± 0.13
PPTL	0.372 ± 0.08	0.378 ± 0.05	0.372 ± 0.06	0.427 ± 0.05	0.413 ± 0.07	0.554 ± 0.17

Table 2. The average accuracy and standard deviation of various networks for each training mode. The partial prior rate is 70% in PPTL.

Results

Results of classification and technical verification

Figure 6 shows the distribution of classification results for 10-fold cross-validation for each subject, and Table 2 shows the classification results and standard deviation for all subjects on each algorithm. According to Fig. 6, the within mode exhibits poor classification performance, with the classification accuracy for most subjects only just exceeding the chance level for the four-class task⁴⁸. Without transfer learning, all network models fail to exhibit superior classification performance. The transfer learning algorithms improve the classification accuracy of each deep learning network by more than 10%, as shown both in the classification results of adaptive transfer learning and in PPTL. The better classification performance is exhibited by both transfer learning algorithms when SACNN is used as the base model. In particular, SACNN-PPTL proposed in this paper exhibits the best classification performance with an accuracy of 55.4% ± 0.17. In addition to the experimental results of deep learning and transfer learning algorithms, the classification result (27.4% ± 0.06) of FBCSP¹⁶ in the within mode has been calculated to facilitate the comparison of traditional machine learning algorithm performance. All algorithms' classification accuracies across the three training modes followed a normal distribution (after excluding a few outliers), meeting the conditions for ANOVA. There is no significant improvement ($P > 0.05$) in classification accuracy in within mode but significant improvement ($P < 0.05$) in both adapt and PPTL when SACNN is used as the base model compared to Deep ConvNets. In SACNN, the classification results of adapt and PPTL are significantly higher than within ($P < 0.05$), while PPTL is significantly higher than adapt ($P < 0.05$).

Table 3 shows the technical validation results of each network within mode. In Table 3, SACNN does not show any advantage over other algorithms within mode. From the F1 Score, Precision and Recall it shows that there is not much difference in the model's classification of the four types of tasks.

Table 4 shows the technical validation results of each network in PPTL. In Table 4, the classification accuracies of all network models based on the PPTL proposed in this paper are improved, indicating the universality of the algorithm. In the results of F1 score, Precision, and recall, SACNN outperforms other algorithms, which indicates that the SACNN proposed in this paper is more compatible with the transfer learning base modeling

Network	F1	Precision	Recall	Accuracy
Deep ConvNets	0.268	0.271	0.271	0.258
EEGNet	0.264	0.270	0.265	0.261
FBCNet	0.285	0.303	0.286	0.286
ATCNet	0.279	0.283	0.281	0.285
EEG Conformer	0.285	0.291	0.286	0.284
SACNN	0.316	0.302	0.302	0.305

Table 3. The results of technical verification for each network in within mode.

Network	F1	Precision	Recall	Accuracy
Deep ConvNets	0.367	0.386	0.366	0.372
EEGNet	0.369	0.379	0.369	0.378
FBCNet	0.368	0.378	0.367	0.381
ATCNet	0.422	0.439	0.419	0.427
EEG Conformer	0.408	0.429	0.405	0.413
SACNN	0.531	0.563	0.541	0.554

Table 4. The results of technical verification for each network in PPTL mode. The partial prior rate is 70% in PPTL.

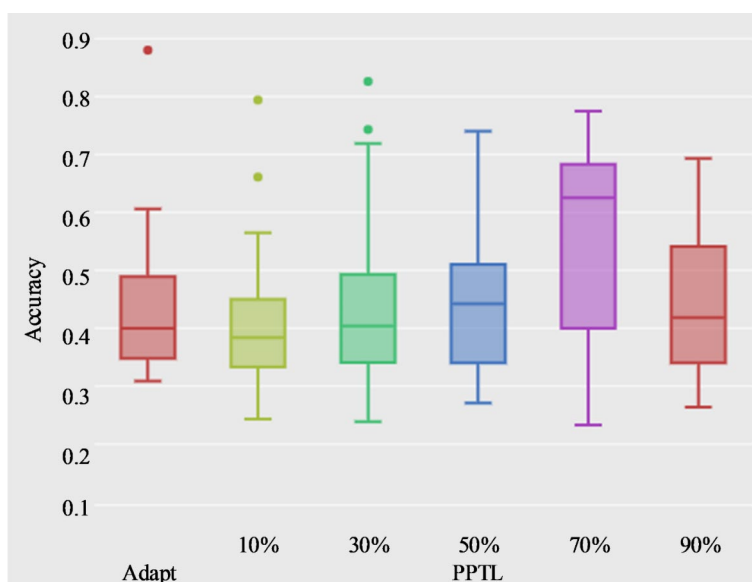


Fig. 7. The classification accuracy of adaptive transfer learning and PPTL with different partial prior rate. The horizontal axis represents the distribution of classification accuracy for all subjects in the SACNN network based on adapt and PPTL modes, while the vertical axis indicates the classification accuracy rate.

for the four types of tasks. Meanwhile, the average results of technical validation are not much different from the classification results, which indicates the reliability of the data and algorithms.

Partial prior transfer learning performance metrics

Figure 7 presents the classification accuracy and F1 scores with different partial prior rates, which indicates that the average classification accuracy increases with more prior data, and the highest classification accuracy appears when the partial prior rate is 70%. The classification results are reliable regardless of the partial prior rate, based on the results of F1 and classification accuracy. The classification accuracies of adapt and PPTL with different partial prior rate models based on SACNN are followed a normal distribution (after excluding a few outliers), meeting the conditions for ANOVA. There is no significant difference ($P > 0.05$) between PPTL with 10% and 30% partial prior rate and SACNN-based adaptive transfer learning. PPTL is significantly ($P < 0.05$) higher than adaptive transfer learning when the partial prior rate is higher than 50%. The classification accuracy

of PPTL is highest when the partial prior rate is 70%. The classification accuracy decreases when the partial prior rate exceeds 70%.

Confusion matrix

The results of the confusion matrix for the classification results of 22 subjects in adaptive transfer learning and PPTL are shown in Fig. 8. The ability to classify the four classes is demonstrated in both algorithms, with the majority of samples in each class being categorized correctly. The task of rotation is classified correctly with more samples but is also the most likely to identify the other classes as it. The differences in the classification of each of the other three types of tasks are small, especially in PPTL. The sample size of the other three types of tasks classified correctly is about 13%, indicating the stability of the dataset and algorithm. It can also be seen that the classification result of PPTL is higher than adaptive transfer learning.

t-SNE feature visualization

The feature vectors of the validation set output by SACNN-PPTL are reduced to a 2D plane by t-SNE⁴⁹. The t-SNE visualization results of subject 1 and subject 11 are selected for convenience of observation. Adaptive transfer learning as a model transfer-based approach is chosen as a comparison to observe the clustering effect of the features output by the algorithms. The t-SNE visualization results are shown in Fig. 9, where the four colored dots represent the four classes of MI tasks. The feature vectors of both algorithms are well clustered according to the visualization results. Although some feature vectors are confounded in the results of subject 1, the clustering centers of each class of tasks are distinguished significantly, and the model is effective in classifying the samples.

Discussion

Based on the experimental results of SACNN-PPTL on four-types MI data from stroke patients, the main contributions of this paper are as follows: (1) SACNN is proposed as the backbone network for EEG decoding and transfer learning. (2) The proposal of PPTL demonstrates improved classification performance in stroke patients across four types of MI tasks. (3) The decoding performance of SACNN-PPTL in the four types of MI tasks for stroke patients is significantly higher than that of other deep learning and transfer learning comparative algorithms ($P < 0.05$). In relation to the main contributions of this paper, the following sections will discuss the model's spatial learning capabilities, the advantages of partial priors, model complexity, and the algorithm's significance for rehabilitation.

The spatial learning ability of the network is conducive to decoding MI tasks for unilateral upper limb

The weak spatial specificity of multi-task for unilateral upper limb increases the difficulty of MI decoding. Lee et al.'s research highlights the importance of spatial features in decoding MI tasks for unilateral upper limb⁵⁰, particularly using classical feature extraction methods like common spatial patterns, which to some extent address the issue of poor spatial specificity in MI decoding for unilateral upper limb. Zhang et al. designed a decoding system for three types of MI tasks for unilateral upper limb, which enhanced the spatial specificity of features through dynamic brain functional networks, providing more effective features for further classification⁵¹. Traditional machine learning techniques can only extract spatial features in a limited number of dimensions, and it is challenging to actively enhance the correlation between channels through algorithms, which limits the model's ability to fully grasp both local and global spatial features of EEG data. In this paper, the temporal feature module retains all channel features during the time feature extraction process, reducing the impact of temporal

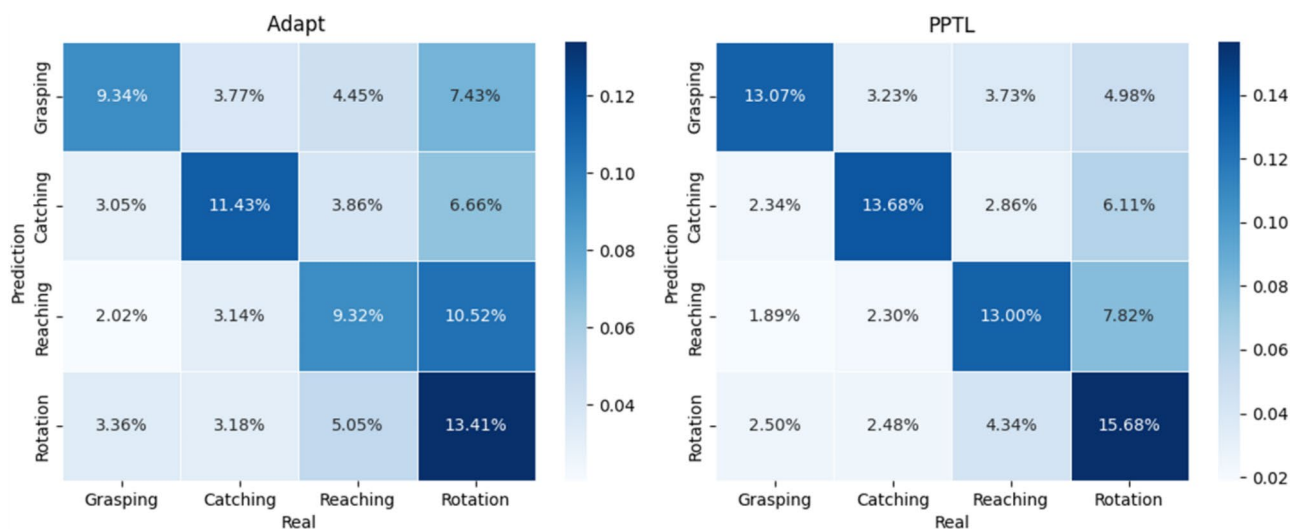


Fig. 8. The results of confusion matrix in adaptive transfer learning and PPTL based on SACNN. The horizontal axis represents the true values, and the vertical axis represents the predicted values.

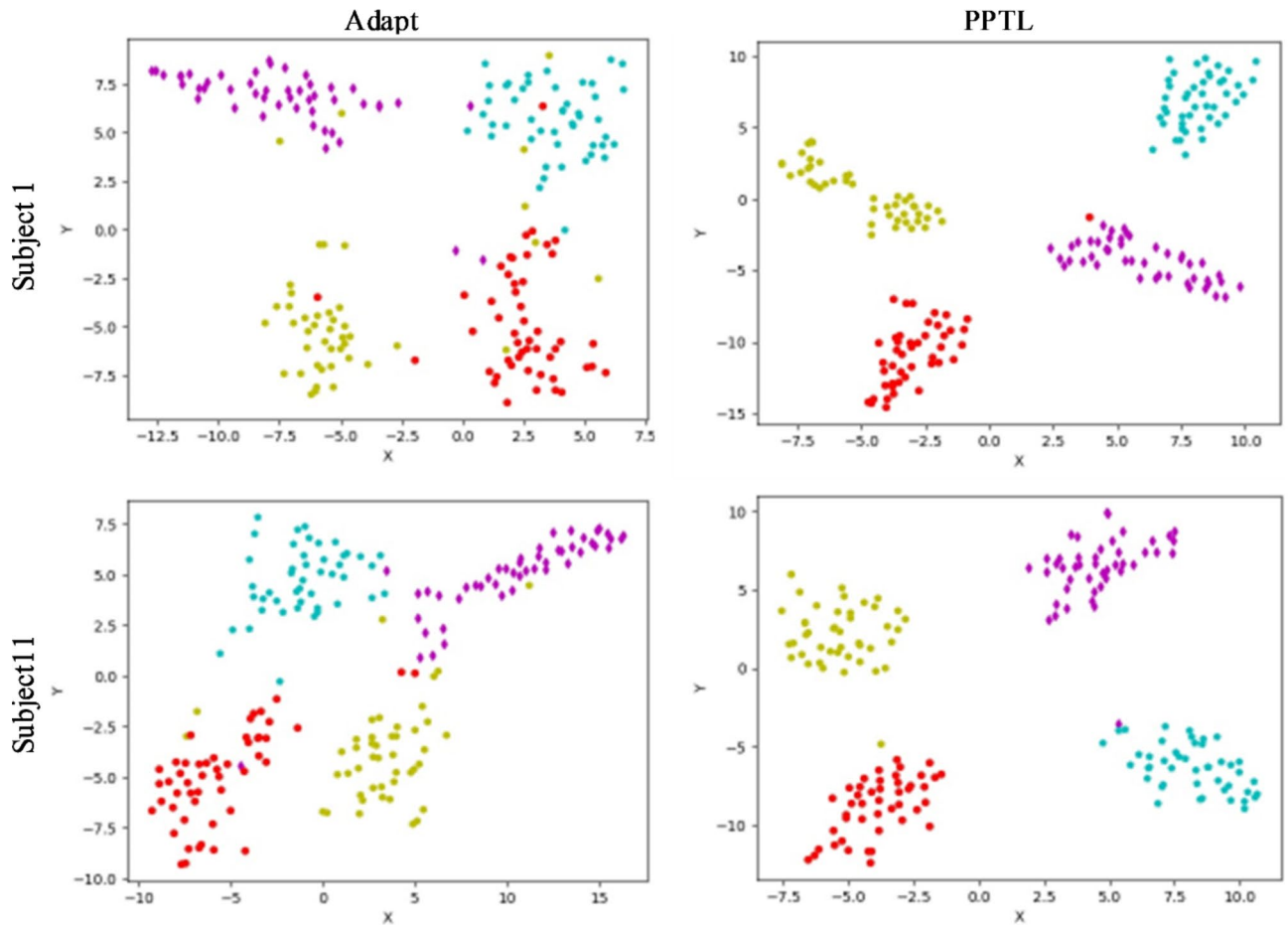


Fig. 9. The feature visualization of t-SNE in adaptive transfer learning and PPTL is based on SACNN. The four different colors of dots represent four types of MI tasks. The horizontal and vertical axes represent the two dimensions after dimensionality reduction using t-SNE.

continuity and enhancing the variation in energy differences within channels. This provides a more spatially distinctive feature map for the subsequent spatial feature extraction module. The spatial feature module further enhances the spatial weights of the feature map through a SA mechanism and spatial convolution, thereby promoting the model's sensitivity to the spatial attributes of EEG signals.

Channel-based self-attention in SACNN is more advantageous for spatial feature extraction, especially excelling in transfer learning. Spatial convolution combined with self-attention feature extraction is typically parallel to temporal feature extraction⁵²⁻⁵³, with both types of features given equal weight. The decision layer classifies based on the combination of these features. Both parallel and serial cascades have been investigated, but the serial cascade of spatiotemporal features demonstrates superior efficacy. In this paper, we use the serial cascade, specifically computing self-attention in the earlier network layers, which better facilitates spatial feature extraction and model transfer. Similar serial cascades⁵⁴ have achieved high classification accuracy in healthy subjects within the BCI IV 2a dataset. However, it is noteworthy that, in comparison to the aforementioned spatial self-attention mechanism, SACNN strategically pre-positions the self-attention layer to preserve as much temporal information as possible. This expansion in input temporal feature dimensions subsequently increases the parameter count in the self-attention layer's linear component, thereby contributing to enhanced computational overhead. Nonetheless, in decoding unilateral limb movements in stroke patients, the pre-positioned spatial self-attention module in the network is more conducive to enhancing spatial features, offering advantages in model prediction and transfer. This pre-positioning method of spatial self-attention may be more sensitive to data with minimal MI feature differences, such as data from stroke patients or unilateral upper limb MI data with less task distinction. Enhancing the network's capacity to learn spatial features via the pre-positioned spatial self-attention module can lead to improved classification accuracy. The classification accuracy of FBCSP is similar to that of the deep learning within mode, indicating that the multi-task decoding of unilateral upper limb in stroke patients is challenging, making it difficult to achieve high classification accuracy using only the subject's own data. However, through model transfer, the decoding accuracy in deep learning demonstrates a significant improvement, which may offer new insights for enhancing the classification accuracy of FBCSP on such challenging datasets.

Partial prior improves the specificity of the target domain

Model-based transfer learning typically uses generalized deep learning models as a foundation, often demonstrating superior classification accuracy in MI decoding²³. When using healthy subjects as the source domain for model transfer to stroke patients, high classification accuracy can also be achieved³⁶. This indicates that model-based transfer learning effectively enhances the classification performance of deep learning models, even with small-sample datasets. However, the generalization of the base model and the specificity of the target domain are uncontrollable. Precisely managing the generalizability of the base model and its specificity to the target domain is a key factor limiting transfer learning models. According to our classification results on adapt and PPTL, partial prior in the source domain is more beneficial for the base model to optimize towards the target domain during training than using only the source domain. The proportion of priors in the base model should not be too much, as excessive reliance on the target domain during basic training can reduce the generalizability of the base model. If the model too quickly aligns with the target domain, then even reducing the learning rate for adaptive training in the target domain cannot achieve the purpose of secondary optimization. The main reason for the poor classification results in within mode is the model's weak ability to recognize the four types of tasks. The poor performance of the within mode is attributed to the small training sample size, which is inadequate for the model to achieve sufficient training and to fully understand the knowledge of multi-task features. As a result, its generalization capability is inferior to that of transfer learning models.

As demonstrated in the experimental results of this paper, the model avoids overfitting when the partial prior rate is set at 70% (126 trials), leading to an effective improvement in the accuracy of the test set. Yet, it's noteworthy that the classification accuracy of 90% PPTL decreases. This decline is attributed to an excessive partial prior rate for the base model, resulting in overfitting and a reduction in the adaptive learning ability of the target model. The mentioned advantages are corroborated by examining the accuracy and F1 score of the partial prior, providing additional evidence for the validity of PPTL.

In transfer learning, although optimizing the source domain model consumes substantial resources, it significantly enhances the performance for data with distinct features. Traditional model-adaptive transfer learning relies on a fixed source domain model^{30,55}. For different target subjects, the feature distribution of the model is hard to completely align with the target domain through adaptive training in the target domain. The optimization of source domain samples and models has also attracted attention from some researchers. For example, selecting source domain training samples at the feature level⁵⁶ employs a priori knowledge, using the target domain to optimize the source domain. However, feature-based optimization methods often lose more EEG information compared to using deep feature optimization as in SACNN. Similar studies include Xu et al.'s⁵⁷ two-level alignment evaluation method, indicating growing attention to the specificity of source domain samples and base models for the target domain. The experimental results of PPTL further validate the effectiveness of optimizing the source domain model using the target domain, providing a new perspective for recognizing EEG data with insignificant spatial features in stroke patients. This approach is expected to be promoted and applied in decoding MI tasks with low feature differentiation.

Complexity and efficiency of model operation

The training efficiency is reduced in modeling independent subjects due to the use of the target domain's training set in PPTL. However, regardless of the training mode, decoding requires individual modeling for each subject. The difference in transfer learning lies in whether it requires building separate base models. Typically, each subject is trained on an independent base model and a target model, to train the model with stronger generalization, on a dataset of less than 30 subjects. Therefore, in decoding performance studies, to achieve higher classification accuracy, one has to opt for reduced training efficiency. Additionally, EEG data is a small sample dataset, and shallow network structures are commonly used, requiring minimal computational resources for training base and target models. Despite slightly higher computational complexity, PPTL mode is more valuable compared to adapt mode, as it offers a 12% improvement in classification accuracy.

In the calculations of the SACNN-PPTL model, due to the need to compute the base model and target model separately for each subject, the computational resources required are also greater. For example, based on the machine performance used in this paper, under consistent training parameters, training SACNN for 200 epochs takes 24 s with each epoch having a sample size of 160 trials (with 20 in the validation set). In contrast, the base model training for PPTL with 3780 samples (420 in the validation set) takes 512 s for 200 epochs. The training time for the PPTL target model is also 24 s, which is consistent with the within mode duration for SACNN. The computational resources used by PPTL are about 21 times greater than those of the within mode, roughly equivalent to the increase in sample size from 21 additional subjects. Each additional subject in the source domain increases the computational resources by an amount equivalent to one fold of the within mode. In MI tasks of unilateral upper limb, even healthy subjects struggle to achieve high classification accuracy, let alone stroke patients³⁶. Therefore, spending considerable computational resources to achieve nearly a 25% performance improvement is meaningful.

While offline systems can pursue higher algorithm classification performance, online systems prioritize the balance between performance and algorithm complexity. In deep learning, enhanced model performance often results in longer processing times, making deployment more challenging compared to traditional machine learning. Additionally, hybrid deep learning combines the feature extraction capabilities of machine learning with the classification abilities of deep learning, offering certain advantages in addressing the complexities of online deployment and algorithm sophistication in BCI systems¹⁹. Due to the limitations of current experimental conditions, the results for SACNN-PPTL are only offline outcomes; however, the model's potential for online deployment is worthy of discussion. In the online training of the BCI rehabilitation system, each EEG trial of the patient is classified by the target model. The EEG is pre-processed and classified on the BCI rehabilitation training client. The time from MI data generation to classification result output is 0.5 s each time, that is, the

system latency of inference phase for each trial is 0.5 s after a 4-second MI task. In real-world applications, the client only performs pre-processing and prediction, while model training is done on the remote server. The client hardware meets the requirements for running PyTorch 1.7.1, and the remote server hardware configuration is consistent with what is specified in the training section of this paper. When the GPU is upgraded to an RTX 4090 on the remote server, the training time for the base model is reduced to 5.6 min and the training time for the target model is 16 s. With the advent of more powerful hardware, the computational time required for the PPTL algorithm is expected to decrease, progressively alleviating the challenge of prolonged model training durations. Even if the online modeling time is long, a low-performance model can be quickly deployed using the within mode, and it can automatically switch to the PPTL model once it is sent from the remote server. Due to the universality of the transmission protocol and data format, the client only performs data upload and classification, regardless of the algorithm used. Therefore, SACNN-PPTL can share the client environment with other algorithms and is fully compatible with existing BCI rehabilitation training systems. The algorithms and datasets are stored on remote servers, allowing for dynamic adjustment of network parameters and sample size during model training. This enables algorithms to be maintained and updated in a dynamic real-time environment.

Researchers can choose the training mode that best fits their research goals and subjects. For instance, when there are fewer task types and greater task distinguishability, the within mode can be used to quickly obtain a high-performance model for deployment in rehabilitation systems. When task difficulty increases and real-time performance is required, the adapt mode can be used. When the task difficulty is higher and there are more task types, researchers should select an appropriate source domain sample size based on time requirements to enhance classification performance while ensuring the timeliness of model training, deploying using the PPTL mode.

Multi-task of unilateral upper limb are more conducive to rehabilitation

Activation of the motor brain regions based on BCI is the most compelling evidence for stroke rehabilitation. According to the brain map (Fig. 1) of the raw task state data in the α and β bands, four types of tasks exhibit ERD/S phenomenon. Due to the spatial differences between different tasks, performing various MI tasks activates different brain regions, providing greater activation compared to a single grasping task during rehabilitation training. This conclusion demonstrates the positive role of the four types of MI tasks for unilateral upper limb stroke motor rehabilitation and theoretically verifies the distinguishability of these tasks.

In a similar study by Ma et al.¹³, the ERD/S phenomenon in the hand and elbow are very significant, similar to the brain map for the catching task shown in Fig. 1. However, the grasping task in this paper does not show significant ERD/S characteristics, the ipsilateral elevation phenomenon is higher than that observed in Ma et al.'s study. This might be related to the patient data used in this paper. Healthy subjects can flexibly perform movements of both the hand and elbow, whereas patients are typically more accustomed to using the elbow joint (as the hand is prone to paralysis), which is why the elbow joint appears closer to healthy subjects on the brain map. The study by Edelman et al.¹¹ utilized fMRI-EEG multimodal fusion analysis to decode four types of MI tasks for unilateral upper limb. Due to the high spatial resolution of fMRI, the ERD/S presented are concentrated in the precentral gyrus on both sides of the brain, resulting in more prominent activation features. These experimental results are similar to our preliminary experiments conducted with fMRI on healthy subjects¹⁰, confirming the effectiveness of the MI tasks for patient motor rehabilitation discussed herein. This also indirectly explains the significant importance of decoding these four types of tasks in post-stroke motor rehabilitation.

Limitation and future work

A few limitations are observed in this study. The computational efficiency of PPTL is lower than that of simple model transfer. When the sample size is large enough, model transfer can achieve the sharing of the same offline base model among different subjects, while PPTL must train each base model online, increasing computational complexity and reducing computational efficiency. This leads to the selection of models in clinical applications only when the classification performance is poor and insufficient to meet rehabilitation needs. The limited sample size in the existing source domain also restricts the algorithm's performance. Due to current experimental conditions, the model has not yet undergone online testing. Additionally, influenced by the experimental paradigm, the results of this experiment have not been discussed in relation to other studies on MI in stroke patients.

In the future, the development of partial prior should move away from target domain optimization in the base model stage. Instead, priors can be trained before target training during the process of base model to target model transfer. The purpose of prior training is to find the optimal solution between the generalization ability of the base model and the specificity to the target domain, while ensuring training efficiency and maximizing the retention of the base model's generalization ability and specificity. This aims to reduce computational complexity and improve classification accuracy.

Conclusion

This study introduces a new method called SACNN-PPTL for classifying four types of MI tasks in stroke patients. Experimental results demonstrate that the proposed method exhibits excellent classification performance. Additionally, the unique advantages of this method in model transfer, computational complexity, and neuroscientific interpretability are examined. Therefore, the introduction of SACNN-PPTL contributes to the development of EEG decoding theory for stroke patients and provides guiding suggestions for the application of unilateral upper limb rehabilitation training for stroke patients.

Data availability

The data sets in this study are available upon reasonable request by the corresponding authors.

Received: 28 July 2024; Accepted: 6 November 2024

Published online: 15 November 2024

References

- Collaborators, G. S. Global, regional, and national burden of stroke and its risk factors, 1990–2019: A systematic analysis for the global burden of disease study 2019. *Lancet Neurol.* **20**, 795 (2021).
- Wang, Y. et al. Chinese stroke report 2019 (1). *Chin. J. Stroke.* **15**, 1037–1043 (2020).
- Avan, A. et al. Socioeconomic status and stroke incidence, prevalence, mortality and worldwide burden: An ecological analysis from the global burden of disease study 2017 (2219). *Neurology.* **94**, 2219 (2020).
- Das, A., Suresh, S. & Sundararajan, N. A discriminative subject-specific spatio-spectral filter selection approach for EEG based motor-imagery task classification. *Expert Syst. Appl.* **64**, 375–384 (2016).
- Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G. & Vaughan, T. M. Brain–computer interfaces for communication and control. *Clin. Neurophysiol.* **113**, 767–791 (2002).
- Leuthardt, E. C., Schalk, G., Wolpaw, J. R., Ojemann, J. G. & Moran, D. W. A brain–computer interface using electrocorticographic signals in humans. *J. Neural Eng.* **1**, 63 (2004).
- Biasiucci, A. et al. Brain-actuated functional electrical stimulation elicits lasting arm motor recovery after stroke. *Nat. Commun.* **9**, 2421 (2018).
- Mane, R., Chouhan, T. & Guan, C. BCI for stroke rehabilitation: Motor and beyond. *J. Neural Eng.* **17**, 041001 (2020).
- Ma, J. et al. Recognizable rehabilitation movements of multiple unilateral upper limb: An fMRI study of motor execution and motor imagery. *J. Neurosci. Methods.* **392**, 109861 (2023).
- Yang, B., Ma, J., Qiu, W., Zhang, J. & Wang, X. The unilateral upper limb classification from fMRI-weighted EEG signals using convolutional neural network. *Biomed. Signal Process. Control.* **78**, 103855 (2022).
- Edelman, B., Baxter, B. & He, B. in *7th International IEEE/EMBS Conference on Neural Engineering (NER)*. 194–197 (IEEE). (2015).
- Edelman, B. J., Baxter, B. & He, B. EEG source imaging enhances the decoding of complex right-hand motor imagery tasks. *IEEE Trans. Biomed. Eng.* **63**, 4–14 (2015).
- Ma, X., Qiu, S., Wei, W., Wang, S. & He, H. Deep channel-correlation network for motor imagery decoding from the same limb. *IEEE Trans. Neural Syst. Rehabil. Eng.* **28**, 297–306 (2019).
- Fadiyah, A. U. & Djamal, E. C. in *6th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*. 28–33 (IEEE, 2019).
- Ramoser, H., Muller-Gerking, J. & Pfurtscheller, G. Optimal spatial filtering of single trial EEG during imagined hand movement. *IEEE Trans. Rehabil. Eng.* **8**, 441–446 (2000).
- Ang, K. K., Chin, Z. Y., Zhang, H. & Guan, C. in *IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence)*. 2390–2397 (IEEE, 2008).
- Moaveninejad, S. et al. Fractal dimension as a discriminative feature for high accuracy classification in motor imagery EEG-based brain-computer interface. *Comput. Methods Programs Biomed.* **244**, 107944 (2024).
- Ferracuti, F., Iarlori, S., Mansour, Z., Monteriù, A. & Porcaro, C. Comparing between different sets of preprocessing, classifiers, and channels selection techniques to optimise motor imagery pattern classification system from EEG pattern recognition. *Brain Sci.* **12**, 57 (2021).
- Alzahab, N. A. et al. Hybrid deep learning (hDL)-based brain-computer interface (BCI) systems: A systematic review. *Brain Sci.* **11**, 75 (2021).
- Zhang, X. et al. in *IEEE International Conference on Pervasive Computing and Communications (PerCom)*. 1–10 (IEEE, 2018).
- Tang, X., Wang, T., Du, Y. & Dai, Y. Motor imagery EEG recognition with KNN-based smooth auto-encoder. *Artif. Intell. Med.* **101**, 101747 (2019).
- Lawhern, V. J. et al. EEGNet: A compact convolutional neural network for EEG-based brain–computer interfaces. *J. Neural Eng.* **15**, 056013 (2018).
- Mane, R., Robinson, N., Vinod, A. P., Lee, S. W. & Guan, C. in *42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. 2950–2953 (IEEE, 2020).
- Schirrmester, R. T. et al. Deep learning with convolutional neural networks for EEG decoding and visualization. *Hum. Brain Mapp.* **38**, 5391–5420 (2017).
- Altaheri, H., Muhammad, G. & Alsulaiman, M. Physics-informed attention temporal convolutional network for EEG-based motor imagery classification. *IEEE Trans. Industr. Inf.* **19**, 2249–2258 (2022).
- Song, Y., Zheng, Q., Liu, B. & Gao, X. EEG conformer: Convolutional transformer for EEG decoding and visualization. *IEEE Trans. Neural Syst. Rehabil. Eng.* **31**, 710–719 (2022).
- Hou, Y. et al. GCNs-net: A graph convolutional neural network approach for decoding time-resolved eeg motor imagery signals. *IEEE Trans. Neural Networks Learn. Syst.* (2022).
- Ju, C. & Guan, C. Tensor-cspnet: A novel geometric deep learning framework for motor imagery classification. *IEEE Trans. Neural Netw. Learn. Syst.* **34**, 10955–10969 (2022).
- Wang, X. et al. A two-branch depthwise CNN for decoding EEG-based multi-class motor imagery. *Neural Netw.* **170**, 312–324 (2024).
- Zhang, K., Robinson, N., Lee, S. W. & Guan, C. Adaptive transfer learning for EEG motor imagery classification with deep convolutional neural network. *Neural Netw.* **136**, 1–10 (2021).
- Mattioli, F., Porcaro, C. & Baldassarre, G. A 1D CNN for high accuracy classification and transfer learning in motor imagery EEG-based brain-computer interface. *J. Neural Eng.* **18**, 066053 (2022).
- Sun, B., Wu, Z., Hu, Y. & Li, T. Golden subject is everyone: A subject transfer neural network for motor imagery-based brain computer interfaces. *Neural Netw.* **151**, 111–120 (2022).
- Peng, Z., Li, H., Zhao, D. & Pan, C. Reducing the dimensionality of SPD matrices with neural networks in BCI. *Mathematics.* **11**, 1570 (2023).
- Zhang, Y. et al. Transfer learning algorithm design for feature transfer problem in motor imagery brain-computer interface. *China Commun.* **19**, 39–46 (2022).
- Peterson, V. et al. Transfer learning based on optimal transport for motor imagery brain–computer interfaces. *IEEE Trans. Biomed. Eng.* **69**, 807–817 (2021).
- Ma, J., Yang, B., Rong, F., Gao, S. & Wang, W. Motor imagery decoding using source optimized transfer learning based on multi-loss fusion CNN. *Cogn. Neurodyn.* 1–14 (2024).
- Naghdi, S., Ansari, N. N., Mansouri, K. & Hasson, S. A neurophysiological and clinical study of Brunnstrom recovery stages in the upper limb following stroke. *Brain Inj.* **24**, 1372–1378 (2010).
- Saini, M. & Satija, U. An effective and robust framework for ocular artifact removal from single-channel EEG signal based on variational mode decomposition. *IEEE Sens. J.* **20**, 369–376 (2019).

39. Mognon, A., Jovicich, J., Bruzzone, L. & Buiatti, M. A. D. J. U. S. T. An automatic EEG artifact detector based on the joint use of spatial and temporal features. *Psychophysiology*. **48**, 229–240 (2011).
40. Yang, K. et al. High gamma band EEG closely related to emotion: Evidence from functional network. *Front. Hum. Neurosci.* **14**, 89 (2020).
41. He, K., Zhang, X., Ren, S. & Sun, J. in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 770–778.
42. Chen, J., Wang, D., Yi, W., Xu, M. & Tan, X. Filter bank sinc-convolutional network with channel self-attention for high performance motor imagery decoding. *J. Neural Eng.* **20**, 026001 (2023).
43. LeCun, Y., Bottou, L., Bengio, Y. & Haffner, P. Gradient-based learning applied to document recognition. *Proc. IEEE*. **86**, 2278–2324 (1998).
44. Corsi, M. C. et al. Integrating EEG and MEG signals to improve motor imagery classification in brain–computer interface. *Int. J. Neural Syst.* **29**, 1850014 (2019).
45. Vaswani, A. et al. Attention is all you need. *Adv. Neural Inf. Process. Syst.* **30** (2017).
46. Ioffe, S. Batch normalization: Accelerating deep network training by reducing internal covariate shift. arXiv preprint arXiv:1502.03167 (2015).
47. Clevert, D. A. Fast and accurate deep network learning by exponential linear units (elus). arXiv preprint arXiv:1511.07289 (2015).
48. Combrisson, E. & Jerbi, K. Exceeding chance level by chance: The caveat of theoretical chance levels in brain signal classification and statistical assessment of decoding accuracy. *J. Neurosci. Methods*. **250**, 126–136 (2015).
49. Van der Maaten, L. & Hinton, G. Visualizing data using t-SNE. *J. Mach. Learn. Res.* **9** (2008).
50. Lee, M., Jeong, J. H., Kim, Y. H. & Lee, S. W. Decoding finger tapping with the affected hand in chronic stroke patients during motor imagery and execution. *IEEE Trans. Neural Syst. Rehabil. Eng.* **29**, 1099–1109 (2021).
51. Zhang, Q. et al. Tensor-based dynamic brain functional network for motor imagery classification. *Biomed. Signal Process. Control*. **69**, 102940 (2021).
52. Liu, X. et al. Parallel spatial–temporal self-attention CNN-based motor imagery classification for BCI. *Front. NeuroSci.* **14**, 587520 (2020).
53. Tao, W. et al. ADFCNN: attention-based dual-scale fusion convolutional neural network for motor imagery brain–computer interface. *IEEE Trans. Neural Syst. Rehabil. Eng.* (2023).
54. Fang, T. et al. Noninvasive neuroimaging and spatial filter transform enable ultra low delay motor imagery EEG decoding. *J. Neural Eng.* **19**, 066034 (2022).
55. Roy, A. M. Adaptive transfer learning-based multiscale feature fused deep convolutional neural network for EEG MI multiclassification in brain–computer interface. *Eng. Appl. Artif. Intell.* **116**, 105347 (2022).
56. Huang, G. et al. Discrepancy between inter- and intra-subject variability in EEG-based motor imagery brain-computer interface: Evidence from multiple perspectives. *Front. NeuroSci.* **17**, 1122661 (2023).
57. Xu, D., Sun, Y. & Li, M. An adaptive cross-class transfer learning framework with two-level alignment. *Biomed. Signal Process. Control*. **86**, 105155 (2023).

Author contributions

J.M.: experiments, methodology, visualization, writing review and editing; W.M.: resources, Investigation; J.Z.: data curation; Y.L.: experiments; B.Y.: Methodology and writing review; C.S.: supervision and resources.

Funding

This study was funded by the National Natural Science Foundation of China (82272612); National High Level Hospital Clinical Research Funding and Elite Medical Professionals Project of China-Japan Friendship Hospital (ZRJY2021-QM02); Young Elite Scientist Sponsorship Program By Bast (BYESS2023173).

Declarations

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/s41598-024-79202-8>.

Correspondence and requests for materials should be addressed to J.M., B.Y. or C.S.

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

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024