TDLNet: Transfer Data Learning Network for Cross-Subject Classification Based on Multiclass Upper Limb Motor Imagery EEG

Jingfeng Bi and Ming Chu

***Abstract*— The limited number of brain-computer interface based on motor imagery (MI-BCI) instruction sets for different movements of single limbs makes it difficult to meet practical application requirements. Therefore, designing a single-limb, multi-category motor imagery (MI) paradigm and effectively decoding it is one of the important research directions in the future development of MI-BCI. Furthermore, one of the major challenges in MI-BCI is the difficulty of classifying brain activity across different individuals. In this article, the transfer data learning network (TDLNet) is proposed to achieve the cross-subject intention recognition for multiclass upper limb motor imagery. In TDLNet, the Transfer Data Module (TDM) is used to process cross-subject electroencephalogram (EEG) signals in groups and then fuse cross-subject channel features through two one-dimensional convolutions. The Residual Attention Mechanism Module (RAMM) assigns weights to each EEG signal channel and dynamically focuses on the EEG signal channels most relevant to a specific task. Additionally, a feature visualization algorithm based on occlusion signal frequency is proposed to qualitatively analyze the proposed TDLNet. The experimental results show that TDLNet achieves the best classification results on two datasets compared to CNN-based reference methods and transfer learning method. In the 6-class scenario, TDLNet obtained an accuracy of 65%±0.05 on the UML6 dataset and 63%±0.06 on the GRAZ dataset. The visualization results demonstrate that the proposed framework can produce distinct classifier patterns for multiple categories of upper limb motor imagery through signals of different frequencies. The ULM6 dataset is available at https://dx.doi.org/10.21227/8qw6-f578.**

***Index Terms*— Brain-computer interface, deep learning, electroencephalogram, motor imagery.**

I. Introduction

BRAIN-COMPUTER interface (BCI) is an emerging discipline field produced at the intersection of brain science and information science, which studies how to establish direct communication and control channels between the brain and external devices to realize the information exchange between the brain and the device [1], [2]. According to the different signal acquisition methods, BCI can be divided into two categories: invasive and non-invasive. Invasive BCI is the surgical implantation of electrode arrays into the skull to directly record or stimulate brain neurons to achieve communication with the outside world, because invasive BCI requires surgery, so it is currently only used in animal experiments or severely paralyzed patients. Compared with invasive BCI, non-invasive BCI is more widely used, but the signal-to-noise of the signal recorded by the non-intrusive acquisition device is relatively low, and researchers continue to study the coding experimental paradigm and decoding methods to improve the application ability of BCI [3]. There have been many algorithms developed for electroencephalogram (EEG) pattern classification in various BCI applications [4], [5], [6], [7]. In non-invasive BCI, researchers have extensively and deeply explored brain signal encoding and decoding methods for application scenarios in different fields, forming active, reactive and passive BCI interaction modes. Active BCI refers to the output of control signals that reflect brain activity, independent of external events, and are typically represented by motor imagery (MI) [8], [9], [10], [11], [12]. Reactive BCI refers to brain activity triggered by external stimuli, and the system then produces output according to the brain's response, which mainly includes steady-state visual evoked potential (SSVEP) BCI [13], [14], [15] and event-related potential (ERP) BCI [16], [17], [18]. Passive BCI does not aim at control, providing the computer with the hidden state of the brain in the process of human-computer interaction, so as to facilitate the computer to make timely adjustments and realize humanized interaction, which mainly includes emotional BCI [19], [20], [21], mental workload BCI [22], [23], [24]. Among these BCI interaction modes, MI used in BCI systems has received increasing attention because it allows users to produce control signals without external stimuli.

In current research on brain-computer interface based on motor imagery (MI-BCI), a novel algorithm, namely temporally constrained sparse group spatial pattern (TSGSP) was proposed in [25] for the simultaneous optimization of filter bands and time window within CSP to further boost classification accuracy of MI-EEG. An experimental study was implemented on three publicly EEG datasets (BCI competition III dataset IIIa, BCI competition IV datasets IIa, and BCI competition IV dataset IIb) to validate the effectiveness of

Manuscript received 22 March 2023; revised 13 September 2023; accepted 6 October 2023. Date of publication 10 October 2023; date of current version 13 October 2023. This work was supported in part by the Beijing Municipal Natural Science Foundation under Grant 3202021 and in part by the National Natural Science Foundation of China under Grant 51875046. *(Corresponding author: Ming Chu.)*

The authors are with the School of Automation, Beijing University of Posts and Telecommunications, Beijing 100876, China (e-mail: bupt_bjf@bupt.edu.cn; chuming_bupt@bupt.edu.cn).

Digital Object Identifier 10.1109/TNSRE.2023.3323509

TSGSP. Sakhavi et al. [26] proposed a classification framework for MI data by introducing a new temporal representation of the data and also utilizing a convolutional neural network (CNN) architecture for classification. The new representation is generated from modifying the filter-bank common spatial patterns method, and the CNN is designed and optimized accordingly for the representation. Tabar and Halici [27] investigated convolutional neural networks (CNN) and stacked autoencoders (SAE) for classifying EEG motor imagery signals. They proposed a new deep network that combines CNN and SAE. In [28], a subject-independent framework based on deep convolutional neural networks (CNNs) was proposed. This framework formulates the discriminative feature representation by combining the spectral-spatial input, embedding the diversity of EEG signals, and a feature representation learned from the CNN through a fusion technique that integrates various discriminative brain signal patterns. While most previous MI-BCI research has produced excellent results, current BCI systems based on MI are typically effective only in distinguishing left and right motor execution/imagery.

The limited number of MI-BCI instruction sets based on different movements of single limbs makes it difficult to meet practical application requirements. In addition, controlling peripherals by imagining the movements of different limbs often leads to inconsistencies between motor intentions and actual peripheral execution instructions, called cognitive disconnection [29], which increases the cognitive load of patients and seriously affects the use of the BCI system. Therefore, designing a single limb multi-category MI paradigm and effectively decoding it is one of the important research directions in the future development of MI-BCI. In recent years, researchers have begun to explore the fine MI paradigm of multiple movements in a single joint and multiple joints in a single limb. A multi-directional convolution neural network-bidirectional long short-term memory network (MDCBN)-based deep learning framework was proposed in [30], for the classification of six directions of arm extension, that achieves a total average correlation coefficient classification effect of 0.45 in motor imagery sessions. A time-distributed attention network (TD-Atten) was proposed in [31] to adaptively assign different weights to different classes and frequency bands of the input multiband Common Spatial Pattern (CSP) features, and obtained the accuracies of 46.8% in the 5-class scenario and 53.4% in the 4-class scenario. A novel fMRI-weighted Convolutional Neural Network (CNN) was designed in [32] to reassign each channel's weight based on brain activation areas to improve classification accuracy, and the average classification accuracy of fMRI-weighted CNN is 47.0%. In the multi-class scenario, the current study has achieved good results in the MI-BCI. However, classifying the brain activity across different individuals remains a challenge for MI-BCI.

Transfer learning (TL) has been proven to be one of the most significant techniques for cross-subject classification in electroencephalogram-based brain-computer interfaces (BCI). Hence, it is widely used to address the challenges of cross-session and cross-subject variability with more accurate intention prediction [33]. Wu et al. [34] proposed a parallel multiscale filter bank convolutional neural network (MSFBCNN) for motor imagery (MI) classification, which is a layered end-to-end network structure. To enhance the transfer learning ability, a network initialization and fine-tuning strategy were proposed to train an individual model for inter-subject classification on small datasets. The transfer learning experiments indicate that the network can build an individual model and obtain acceptable results in inter-subject classification. In [35], a new deep learning (DL) architecture for motor imagery (MI) based brain-computer interfaces (BCIs) called EEGSym was presented, which aims to improve previous state-of-the-art performances on MI classification by overcoming inter-subject variability and reducing BCI inefficiency. Hu et al. [36] proposed a multi-scale adaptive transformer network (MSATNet) for motor imagery classification. Therein, a multi-scale feature extraction (MSFE) module to extract multi-band highly-discriminative features was designed. Through the adaptive temporal transformer (ATT) module, the temporal decoder and multi-head attention unit were used to adaptively extract temporal dependencies. Efficient transfer learning was achieved by fine-tuning target subject data through the subject adapter (SA) module. Zhao et al. [37] proposed a convolutional neural network (CNN) with an end-to-end serial-parallel (SP) structure followed by transfer learning. The serial module was used to extract the rough features in time-frequency-space domain, and the parallel module was used for fine feature learning at different scales. Meanwhile, a freeze-and-retrain fine-tuning transfer learning strategy was proposed to improve the cross-subject accuracy.

In this article, the Transfer Data Learning network (TDLNet) is proposed to achieve cross-subject intention recognition for multiclass upper limb motor imagery (elbow flexion, elbow extension, forearm supination, forearm pronation, hand open, and hand close). TDLNet is mainly composed of three key components: the Transfer Data Module (TDM), Inception Module, and Residual Attention Mechanism Module (RAMM). The TDM performs migration and expansion processing on the data of the two subjects and concatenates the data in the depth direction according to categories. The Inception Module captures multiscale time information from input feature maps by applying multiple parallel branches. The RAMM allocates attention weights to different EEG signal channels for different motor imagery tasks, focusing on the most relevant brain signals to improve the accuracy of BCI decoding. In addition, a feature visualization algorithm based on occlusion signal frequency is proposed to qualitatively analyze TDLNet. In summary, the main contributions of our work are listed as follows.

1) The Transfer Data Learning Network (TDLNet) is proposed to achieve the cross-subject intention recognition for multiclass upper limb motor imagery.
2) In TDLNet, the RAMM is proposed for assigning weights to different EEG signal channels for different motor imagery tasks.
3) A feature visualization algorithm based on occlusion signal frequency is proposed to qualitatively analyze the MI-BCI framework.

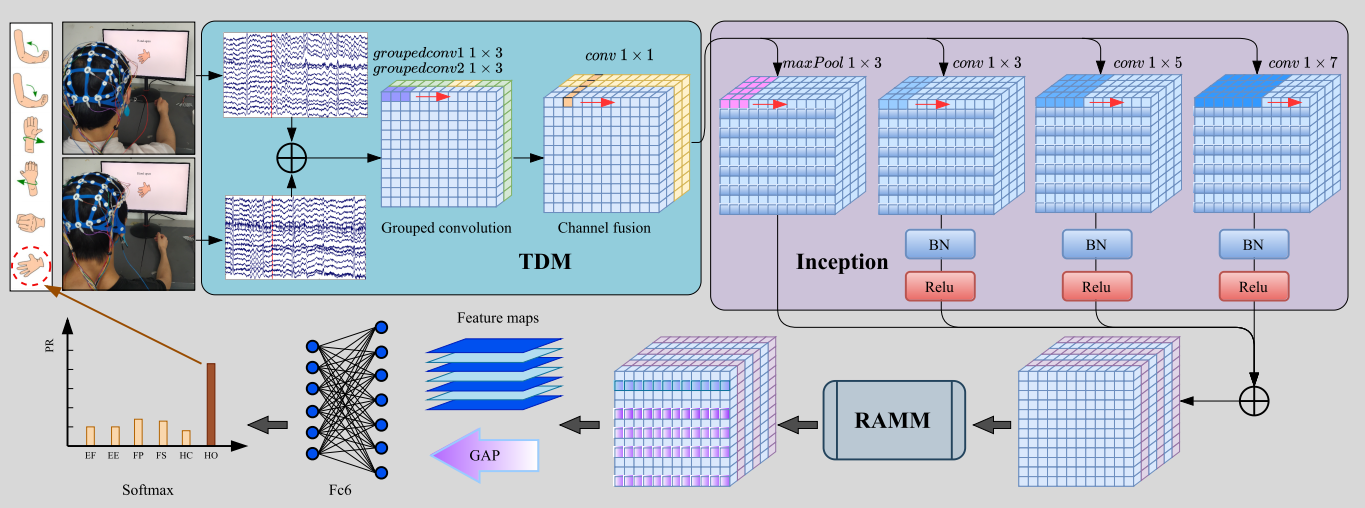

Fig. 1. Architecture of TDLNet, which mainly consists of three components: (1) TDM, (2) Inception Module, and (3) RAMM.

The remainder of this article is organized as follows. In Section II, we provide the architecture of the TDLNet model in detail and describe the proposed feature visualization algorithm based on occlusion signal frequency. In Section III, we show the datasets and implementation details, as well as the ablation studies. TDLNet is experimentally compared to CNN-based methods (DeepConvNet, ShallowConvNet [38] and EEGNet-4,2, EEGNet-8,2 [39]) and transfer learning methods (MSFBCNN [34], EEGSym [35], MSATNet [36]). Three visualization experiments are conducted. In Section IV, we discuss the experimental results and qualitatively analyze the proposed TDLNet using the feature visualization algorithm to demonstrate its effectiveness in cross-subject classification. Finally, Section V concludes this article.

II. Methodology

In this section, Transfer Data Learning Network (TDLNet) based on transfer data and residual attention mechanism is proposed. The key components, Transfer Data Module (TDM), Inception Module, and Residual Attention Mechanism Module (RAMM) are then described in detail. Following this, the feature visualization algorithm based on occluded input signal frequency is proposed for qualitative analysis of TDLNet.

A. Transfer Data Learning Network

The proposed Transfer Data Learning Network (TDLNet) is mainly composed of the Transfer Data Module (TDM), Inception Module, and Residual Attention Mechanism Module (RAMM), as shown in Fig. 1. Concretely, the input of TDM is obtained by splicing the EEG signals of two subjects in the depth direction according to the category. Two layers of grouped convolution are used for feature extraction in the time dimension separately. Then, two layers of one-dimensional convolution fuse the EEG features of the two subjects in the direction of deep channels. The feature map obtained by this process is used as input to the Inception Module. The Inception Module captures multi-scale time information from the input feature maps through multiple parallel branches. Each branch uses convolution kernels of different sizes to extract time features of different scales. Then, multiple branches of different scale features are connected in the depth direction to form the output of the Inception Module. The Residual Attention Mechanism Module assigns weights to each EEG signal channel, indicating its importance to the current task. These weights are learned by the network and used for calculations, effectively highlighting the most relevant parts and suppressing the less important parts. Finally, the output of RAMM is pooled through the global average pooling layer (GAP) and connected to the fully connected connection layer and the softmax layer.

In the training process of TDLNet, the input is denoted as $D = \{(X^1, y^1), \ldots, (X^N, y^N)\}$, where N represents the total number of trials, $X^j \in R^X$ represents the EEG signal matrix of the jth trial. The corresponding class label of trial j is denoted by y^j. It takes values from a set of K class labels L that correspond to the imagined movements performed in each trial, for example, $\forall y^j \in L = \{l_1 = \text{“elbow flexion”}, l_2 = \text{“elbow extension”}, l_3 = \text{“forearm supination”}, l_4 = \text{“forearm pronation”}, l_5 = \text{“hand open”}, l_6 = \text{“hand close”}\}$. Classifier $f(X^j, \omega)$ is trained on these existing trials such that it is able to assign the correct label to new unseen trials. Concretely, we aim to train the classifier to assign the label y^j to X^j using the output of a parametric classifier $f(X^j, \omega) : R^X \rightarrow L$ with parameters ω. The standard machine learning classification method is different from TDLNet in the framework of classification. The standard machine learning classification method includes independent feature extraction and feature classification while TDLNet learn both stages jointly. To better introduce TDLNet, we assume that the classifier $f(X^j, \omega)$ is represented by a standard machine learning pipeline decomposed into two parts: the first part is feature extractor $\phi(X^j, \omega_\phi)$ with parameters ω_ϕ which is learned from the data, and the second part is a classifier g with parameters ω_g which is trained using these features, that is

$$f(X^j, \omega) = g(\phi(X^j, \omega_\phi), \omega_g) \quad (\omega_\phi, \omega_g) \in \omega \tag{1}$$

TABLE I
TDM ARCHITECTURE

Layer	Filters	Sizes	stride	Output
Input				$(C, T, 2)$
Grouped Conv1	$2*4$	$(1,3)$	$(1,2)$	$(C, T/2, 8)$
Batch Normalization				$(C, T/2, 8)$
Relu				$(C, T/2, 8)$
Grouped Conv2	$2*8$	$(1,3)$	$(1,2)$	$(C, T/4, 16)$
Batch Normalization				$(C, T/4, 16)$
Relu				$(C, T/4, 16)$
Conv3	16	$(1,1)$	$(1,1)$	$(C, T/4, 16)$
Batch Normalization				$(C, T/4, 16)$
Conv4	32	$(1,1)$	$(1,2)$	$(C, T/8, 32)$
Batch Normalization				$(C, T/8, 32)$
Relu				$(C, T/8, 32)$

In the process of TDLNet training, by iteratively assigning the value of parameter ω, high probability is assigned to the correct label to minimize the sum of loss functions

$$\omega = \arg\min \sum_{j=1}^{N} loss(y^j, p(l_k|X^j)) \tag{2}$$

where $p(l_k|X^j)$ represents the conditional probability distribution. The loss function is

$$loss(y^j, p(l_k|X^j)) = \sum_{k=1}^{K} -\log(p(l_k|X^j)) \cdot p(y^j = l_k) \tag{3}$$

where $p(y^j = l_k)$ represents the probability that the y^j is l_k.

1) Transfer Data Module: The structure of Transfer Data Module (TDM) is shown in TABLE I, where C is the number of channels, T is the number of time points. The input of TDM is obtained by splicing the EEG signals of two subjects in the depth direction according to the category, which is expressed as

$$X^j = [X_1^j; X_2^j] \tag{4}$$

where X_1^j, X_2^j are EEG signals of the first and second subjects in trial j. Two grouped convolutions are used to perform feature extraction on the input data in the time dimension, which is represented as

$$\begin{bmatrix} \boldsymbol{Z}_{11} \\ \boldsymbol{Z}_{12} \end{bmatrix} = \sigma \begin{bmatrix} groupedConv_{11}(\boldsymbol{W}_{11}, X_1^j) + \boldsymbol{b}_{11} \\ groupedConv_{12}(\boldsymbol{W}_{12}, X_2^j) + \boldsymbol{b}_{12} \end{bmatrix} \tag{5}$$

$$\begin{bmatrix} \boldsymbol{Z}_{21} \\ \boldsymbol{Z}_{22} \end{bmatrix} = \sigma \begin{bmatrix} groupedConv_{21}(\boldsymbol{W}_{21}, \boldsymbol{Z}_{11}) + \boldsymbol{b}_{21} \\ groupedConv_{22}(\boldsymbol{W}_{22}, \boldsymbol{Z}_{12}) + \boldsymbol{b}_{22} \end{bmatrix} \tag{6}$$

where ($\boldsymbol{Z}_{11}$,$\boldsymbol{Z}_{12}$) and ($\boldsymbol{Z}_{21}$,$\boldsymbol{Z}_{22}$) are the features of the two subjects after the *Grouped Conv1* and *Grouped Conv2*, respectively. σ denotes the ReLu. $groupedConv_{11}$ and $groupedConv_{12}$ are two sets of convolution of *Grouped Conv1*. ($\boldsymbol{W}_{11}$,$\boldsymbol{W}_{12}$) and ($\boldsymbol{b}_{11}$,$\boldsymbol{b}_{12}$) are the weights and bias of *Grouped Conv1*, respectively. Two grouped convolutions increase the level of receptive field of the model. While not greatly increasing the network parameters, the optimal weight solution of convolution is increased, which increases the possibility of finding the optimal solution and improves the convergence speed of the network. Then, two layers of one-dimensional convolution fuse the EEG features of the two subjects in the direction of deep channels, which is represented as

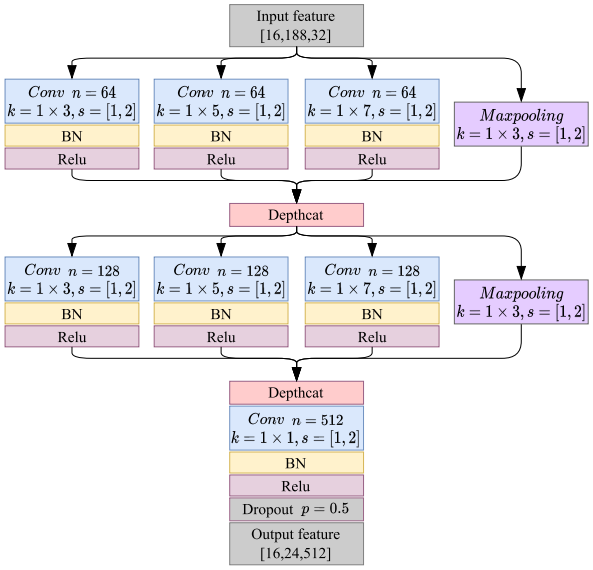

Fig. 2. Structure of TDM.

$$\boldsymbol{Z}_3 = Conv_3(\boldsymbol{W}_3, [\boldsymbol{Z}_{21}; \boldsymbol{Z}_{22}]) + \boldsymbol{b}_3 \tag{7}$$

$$\boldsymbol{Z}_4 = Conv_4(\boldsymbol{W}_4, \boldsymbol{Z}_3) + \boldsymbol{b}_4 \tag{8}$$

The convolution kernel size in this process is 1×1 . *Conv3* focuses on fusing the features of two subjects. *Conv4* extracts deeper temporal features and reduces the size of feature map while further fusing features. Finally, the output $\boldsymbol{TDM}_{\text{out}}$ is obtained after activating feature $\boldsymbol{Z}_4$.

2) Inception Module: The structure of the Inception Module is shown in Fig. 2. Multiple parallel convolution branches with kernel size of $1 \times 3, 1 \times 5, 1 \times 7$ and $Maxpooling$ is used to extract multi-scale time information from the input feature maps. Different scale features are connected in the depth direction.

$$\begin{bmatrix} \boldsymbol{Z}_{\text{I}}^{1\times3} \\ \boldsymbol{Z}_{\text{I}}^{1\times5} \\ \boldsymbol{Z}_{\text{I}}^{1\times7} \\ \boldsymbol{Z}_{\text{I}}^{Max} \end{bmatrix} = \begin{bmatrix} \sigma(Conv(\boldsymbol{W}^{1\times3}, \boldsymbol{TDM}_{\text{out}})) \\ \sigma(Conv(\boldsymbol{W}^{1\times5}, \boldsymbol{TDM}_{\text{out}})) \\ \sigma(Conv(\boldsymbol{W}^{1\times7}, \boldsymbol{TDM}_{\text{out}})) \\ Maxpooling(\boldsymbol{TDM}_{\text{out}}) \end{bmatrix} \tag{9}$$

$$\boldsymbol{Z}_{\text{I}} = \left[\boldsymbol{Z}_{\text{I}}^{1\times3}; \boldsymbol{Z}_{\text{I}}^{1\times5}; \boldsymbol{Z}_{\text{I}}^{1\times7}; \boldsymbol{Z}_{\text{I}}^{Max}\right] \tag{10}$$

where $\boldsymbol{Z}_{\text{I}}$ is the feature extracted from inception block, which is composed of features $\boldsymbol{Z}_{\text{I}}^{1\times3}$, $\boldsymbol{Z}_{\text{I}}^{1\times5}$, $\boldsymbol{Z}_{\text{I}}^{1\times7}$, and $\boldsymbol{Z}_{\text{I}}^{Max}$ of multiple scales. $\boldsymbol{W}^{1\times3}$, $\boldsymbol{W}^{1\times5}$, $\boldsymbol{W}^{1\times7}$ are convolution weights of different scales of inception block. The pointwise convolution with kernel size of 1×1 is used to fuse the features extracted through two inception blocks. The output features $\boldsymbol{F}$ of Inception Module is obtained after passing through *Dropout* layer with probability of 0.5.

3) Residual Attention Mechanism Module: The structure of Residual Attention Mechanism Module (RAMM) is shown in Fig. 3. In RAMM, each channel data of input $\boldsymbol{F}$ is processed separately, and $\boldsymbol{f}_i$ represents the signals of the ith EEG channel. First, the input $\boldsymbol{F}$ is pooled by *Maxpooling* and

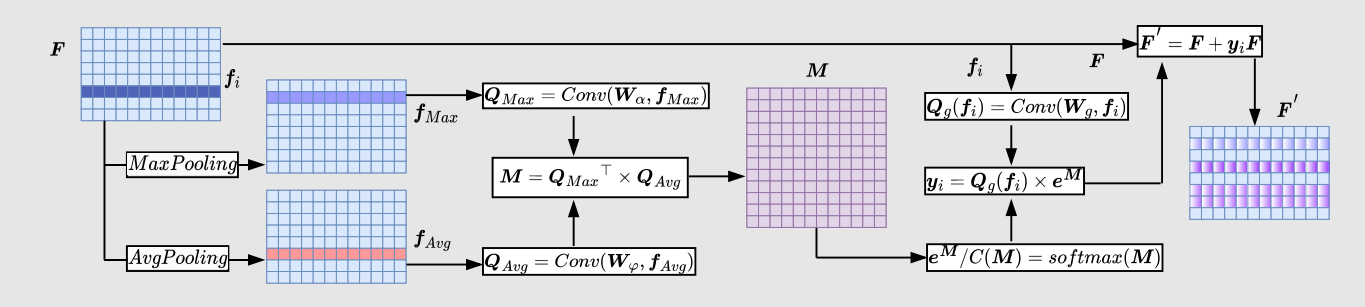

Fig. 3. Structure of RAMM.

Averagepooling, which is expressed as

$$f_{Max} = MaxPooling(F) \tag{11}$$
$$f_{Avg} = AvgPooling(F) \tag{12}$$

where f_{Max} is *Maxpooling* features, f_{Avg} is *Averagepooling* features. The feature maps obtained by the convolution f_{Max} and f_{Avg} are multiplied to get attention weight matrix M, which is expressed as

$$Q_{Max} = Conv(W_\alpha, f_{Max}) \tag{13}$$
$$Q_{Avg} = Conv(W_\varphi, f_{Avg}) \tag{14}$$
$$M = Q_{Max}{}^{\top} \times Q_{Avg} \tag{15}$$

where Q_{Max} and Q_{Avg} are the maximum and average deep features. W_α and W_φ are weights obtained through learning. After that, f_i is convoluted to obtain the features $Q_g(f_i)$ of EEG channel, which is expressed as

$$Q_g(f_i) = Conv(W_g, f_i) \tag{16}$$

The M after *softmax* operation is multiplied by $Q_g(f_i)$ to get the EEG channel attention factor y_i, which is expressed as

$$e^M / C(M) = softmax(M) \tag{17}$$
$$y_i = Q_g(f_i) \times e^M \tag{18}$$

where $C(M)$ is the normalization coefficient. Finally, the EEG channel attention factor y_i is introduced into the input F through the residual connection to obtain the output feature F' with reassigned channel weights

$$F' = F + y_i F \tag{19}$$

Through the above operations, the RAMM assigns weights to each EEG signal channel. The weights W_α and W_φ of each F are fixed, while the weights W_g varies with the channels.

B. Feature Visualization Algorithm

In order to qualitatively analyze TDLNet, a feature visualization algorithm based on occlusion signal frequency is proposed, as shown in Algorithm 1. The test datasets is denoted as $T = \{X^1, \ldots, X^M\}$. Real label of test datasets is denoted as $Y = \{y^1, \ldots y^M\}$, where M represents the total number of test trials. $f(X, \omega)$ is a well-trained TDLNet classifier, where ω is parameters of classifier. First, the test datasets T is input into the classifier to get the prediction label Y_p.

$$Y_p = f(T, \omega) \tag{20}$$

Compare the prediction label Y_p with the real label Y to get the correctly recognized test datasets T_c. Then the T_c is filtered by filters with frequency ranges of ($\delta : 0.5 \sim 3$ HZ, $\theta : 3 \sim 7$ HZ, $\alpha : 7 \sim 13$ HZ, $\beta : 13 \sim 200$ HZ), which is expressed as

$$\begin{cases} T^\delta = filter(T_C, \delta) \\ T^\theta = filter(T_C, \theta) \\ T^\alpha = filter(T_C, \alpha) \\ T^\beta = filter(T_C, \beta) \end{cases} \tag{21}$$

where $T^\delta, T^\theta, T^\alpha, T^\beta$ are the correctly recognized test datasets in different frequency ranges. The test datasets in

Algorithm 1 Feature Visualization Algorithm Based on Occlusion Signal Frequency

Input: Test datasets $T = \{X^1, \ldots, X^M\}$, real label of test datasets $Y = \{y^1, \ldots y^M\}$, well-trained TDLNet classifier $f(X, \omega)$ with parameters ω, feature extraction function $\phi(X^j, \omega_\phi)$ with parameters ω_ϕ.
Output: Scalp topographic maps for different categories and different frequencies.

Step 1: Use the well-trained TDLNet classifier $f(X, \omega)$ with parameters ω to predict the label Y_p for the test datasets T.

$$Y_p = f(T, \omega)$$

Step 2: Compare the predicted label Y_p with the real label Y to get the correctly recognized test datasets T_c.

Step 3: Filter the correctly recognized test datasets T_c using the filters with frequency ranges ($\delta : 0.5 \sim 3$ HZ, $\theta : 3 \sim 7$ HZ, $\alpha : 7 \sim 13$ HZ, $\beta : 13 \sim 200$ HZ) to obtain the filtered datasets T^δ, T^θ, T^α, and T^β using (21).

Step 4: Extract the activated features for each filtered datasets T^δ, T^θ, T^α, and T^β using the feature extraction function $\phi(X^j, \omega_\phi)$ with parameters ω_ϕ to obtain the feature maps F^δ, F^θ, F^α, and F^β using (22).

Step 5: Average the feature maps F^δ, F^θ, F^α, and F^β according to different categories.

Step 6: Draw scalp topographic maps for different categories and different frequencies using the averaged feature maps from Step 5.

different frequency ranges are input into the feature extraction function $\phi(X^j, \omega_\phi)$ of TDLNet to get the activate features of different frequency ranges, which is expressed as

$$\begin{cases} F^\delta = \phi(T^\delta, \omega_\phi) \\ F^\theta = \phi(T^\theta, \omega_\phi) \\ F^\alpha = \phi(T^\alpha, \omega_\phi) \\ F^\beta = \phi(T^\beta, \omega_\phi) \end{cases} \tag{22}$$

The activate features of different frequency ranges are averaged according to different categories. Finally, all scalp topographic maps are drawn by different categories and different frequencies.

III. Experiments and Results

In this section, first, the datasets used in this study and experimental implementation are introduced. After that, ablation studies were performed to verify the effectiveness of each part of TDLNet. Then, the TDLNet is compared with the CNN-based reference methods (DeepConvNet, ShallowConvNet [38] and EEGNet-4,2, EEGNet-8,2 [39]) and transfer learning methods (MSFBCNN [34], EEGSym [35], MSATNet [36]) on two datasets. Finally, TDM is transplanted to the CNN-based reference methods to demonstrate its effectiveness. In addition, experiments related to feature visualization are carried out.

A. Datasets

1) ULM6 Dataset: The ULM6 dataset is collected through our experiments. Ten healthy participants, aged between 24 and 38 years, with an average age of 30 years (standard deviation of 5 years), were recruited for our experiment. Five participants were male, and all the participants were right-handers. The study was conducted in accordance with the Declaration of Helsinki, and subjects provided informed consent. Subjects had normal or corrected-to-normal vision and no history of neurological or psychiatric disorders. The subjects performed six categories of motor imagery upper limb movements, including elbow flexion, elbow extension, forearm supination, forearm pronation, hand open, and hand close. EEG signals were recorded using 16 active Ag/AgCl electrodes with a 16-channel amplifier (OpenBCI CytonDaisy 16-channel Biosensing Board). The sampling frequency was 500 HZ, with the reference electrode on the left earlobe and the ground on the right earlobe. The EEG electrodes were placed according to the international standard 10-20 electrode system. The ULM6 dataset contains 18000 trials (300 trials × 6 categories × 10 subjects). The ULM6 dataset has been uploaded to IEEE DataPort, and is available at "https://dx.doi.org/10.21227/8qw6-f578".

2) GRAZ Dataset: The GRAZ dataset is provided by Ofner et al. [40], which is available from the BNCI Horizon 2020 database at "http://bnci-horizon-2020.eu/database/datasets". The GRAZ dataset consists of electroencephalography (EEG) data from 15 healthy subjects aged between 22 to 40 years, with a mean age of 27 years (standard deviation 5 years). EEG were measured from 61 channels covering frontal, central, parietal and temporal areas using active electrodes and four 16-channel amplifiers (g.tec medical engineering GmbH, Austria). A total of 5400 trials (60 trials × 6 categories × 15 subjects) of six categories were selected from the GRAZ dataset for this study.

TABLE II
ADAM Optimizer Parameters

Parameters	Value	Parameters	Value
Initial learn rate	0.001	Learn rate drop factor	0.1
Gradient threshold	1	Learn rate drop period	200
Max epochs	500	Squared gradient decay factor	0.999
Mini batch size	64	Gradient Decay Factor	0.9

TABLE III
Classification Accuracy (%) of Cross-Subject Ablation Experiments

Methods	Datasets	
	ULM6	GRAZ
TDLNet-w/o-TDM	18%±0.21	17%±0.27
TDLNet-w/o-Inception	25%±0.19	28%±0.19
TDLNet-w/o-RAMM	23%±0.21	24%±0.16
TDLNet	**65%±0.05**	**63%±0.06**

B. Implementation Details

During the cross-subject training of TDLNet, the training data for each experiment consist of 70% of the data from two subjects, and the remaining 30% used as the test datasets. both the ULM6 dataset and the GRAZ dataset consist of 10 and 15 subjects, respectively. Therefore, the number of classification experiments on the two datasets are C_{10}^2 and C_{15}^2, and the experimental results are averaged. The ADAM [41] optimizer was used to train the model, and optimizer parameters are shown in TABLE II. TDLNet was developed using MATLAB R2020b (The MathWorks, Inc., Natick, MA, USA) and was trained on high performance GPU (GeForce RTX 5000) installed on an Intel (R) Core (TM) i7-7000K CPU processor with 64 GB RAM.

In the comparisons with CNN-based reference methods (DeepConvNet, ShallowConvNet [38] and EEGNet-4,2, EEGNet-8,2 [39]) and transfer learning methods (MSFBCNN [34], EEGSym [35], MSATNet [36]), these models were implemented in Tensorflow and Keras following the descriptions found in the corresponding paper. Since their architectures were originally designed for 128 HZ and 250 HZ EEG signals, the EEG signals of ULM6 and GRAZ were down-sampled to the corresponding frequency to suit their structure. These models were trained in the same way as the TDLNet model.

C. Ablation Studies

To evaluate the contribution of each component in TDLNet to the accuracy of cross-subject classification, ablation experiments are conducted. First, the TDM is removed from TDLNet and denoted as TDLNet-w/o-TDM. In the training and classification of TDLNet-w/o-TDM, the EEG data of the two subjects are treated independently. Similarly, the Inception and RAMM components are removed from TDLNet,

TABLE IV
COMPARISONS OF CROSS-SUBJECT CLASSIFICATION ACCURACY (%) WITH CNN-BASED AND TRANSFER LEARNING METHODS

Methods	Datasets	
	ULM6	GRAZ
DeepConvNet [38]	40%±0.17	41%±0.15
ShallowConvNet [38]	35%±0.22	37%±0.16
EEGNet-4,2 [39]	45%±0.17	41%±0.18
EEGNet-8,2 [39]	48%±0.16	50%±0.18
MSFBCNN [34]	45%±0.17	41%±0.19
EEGSym [35]	51%±0.08	50%±0.09
MSATNet [36]	46%±0.14	44%±0.13
TDLNet	**65%±0.05**	**63%±0.06**

TABLE V
ACCURACY (%) OF CROSS-SUBJECT CLASSIFICATION FOR TDM COMBINED WITH CNN-BASED METHODS

Methods	Datasets	
	ULM6	GRAZ
DeepConvNet-TDM	45%±0.15	44%±0.11
ShallowConvNet-TDM	34%±0.16	37%±0.09
EEGNet-4,2-TDM	49%±0.11	46%±0.16
EEGNet-8,2-TDM	53%±0.12	54%±0.11

resulting in models denoted as TDLNet-w/o-Inception and TDLNet-w/o-RAMM, respectively. The classification accuracy of ablation experiments on ULM6 and GRAZ is shown in TABLE III. The TABLE III shows that, compared with TDLNet, the cross-subject classification accuracy of TDLNet-w/o-TDM is significantly reduced by 47% on ULM6 and 46% on GRAZ. Similarly, compared with TDLNet, the cross-subject classification accuracy of TDLNet-w/o-Inception and TDLNet-w/o-RAMM is also significantly reduced, especially in the case of TDLNet-w/o-RAMM, which exhibits a decrease of 42% (65% vs. 23%) on ULM6 and 39% (63% vs. 24%) on GRAZ. The results of ablation experiments indicate that all three components of TDLNet contribute significantly to the cross-subject classification accuracy.

D. Comparisons With CNN-Based Reference and Transfer Learning Methods

The cross-subject performance of TDLNet is compared with the CNN-based reference methods (DeepConvNet, ShallowConvNet [38] and EEGNet-4,2, EEGNet-8,2 [39]) and transfer learning methods (MSFBCNN [34], EEGSym [35], MSATNet [36]) on both the ULM6 and GRAZ datasets. The results of the cross-subject classification experiments are presented in TABLE IV, and ROC curves are depicted in Fig. 4 for comparing TDLNet's performance with other CNN-based reference methods. Compared to ShallowConvNet, the accuracy of TDLNet is improved by 30% on ULM6 and 26% on GRAZ. The cross-subject classification accuracy of TDLNet for six categories achieved 65%±0.05 on ULM6 and 63%±0.06 on GRAZ. Furthermore, the transfer learning method EEGSym attains a classification accuracy of 51%±0.08 on ULM6 and 50%±0.09 on GRAZ, surpassing all CNN-based reference methods discussed in this article. It can be seen from TABLE IV and Fig. 4 that TDLNet outperforms not only other CNN-based reference methods but also transfer learning methods in cross-subject classification on both datasets. These results highlight TDLNet's excellent performance in classifying six categories of upper limb motor imagery across different subjects.

E. TDM Combined With CNN-Based Methods

To assess the effectiveness of the proposed TDM in enhancing other CNN-based methods, we integrated TDM into DeepConvNet, ShallowConvNet [38], and EEGNet-4,2, EEGNet-8,2 [39]. The cross-subject classification accuracies are presented in TABLE V. Comparing the results from TABLE IV and TABLE V, it is evident that, with the exception of ShallowConvNet-TDM, which experienced a slight decrease in performance on ULM6, the classification accuracy of the other CNN-based methods improved to varying degrees when TDM was integrated. This observation underscores the enhanced effect of TDM on cross-subject motor imagery classification tasks in various CNN-based methods.

F. Visualization Experiments

1) Visualization of EEG Source Estimation: The TDLNet proposed in this article is a brain-computer interface (BCI) model based on motor imagery (MI). The principle behind the MI-BCI system is that when a person imagines a movement, specific regions of their brain become activated, leading to changes in their EEG signals. LORETA [42] is used to visualize the source estimation of EEG data for the two datasets used in this article. This source estimation reflects the contributions of multiple sources to scalp EEG signals in a single cortical map. The Fig. 5 shows the EEG signal source estimation for the same action in both datasets, with a time interval of 250 milliseconds between −0.5s and 1s. Fig. 5(a) corresponds to ULM6, while Fig. 5(b) corresponds to GRAZ. This visualization is independent of TDLNet. The routines from the toolbox [43] were employed to calculate the inverse solutions in this visualization. The toolbox is open-source and available for free download at "https://github.com/aojeda/headModel". As demonstrated in Fig. 5, specific areas of the cerebral cortex become activated during motor imagination, resulting in corresponding changes in EEG signals.

2) Feature Visualization Based on Occlusion Signal Frequency: To investigate how TDLNet can successfully decode information from EEG signals, Algorithm 1 is used to visualize the features extracted from the TDLNet, and the results are presented in Fig. 6. The red circles in the figure indicates distinct classifier patterns that can be utilized for differentiation. It can be seen from Fig. 6 that the movements *hand open* and *hand close* exhibit distinct classifier patterns in the frequency ranges $\delta : 0.5 \sim 3\,\text{HZ}$ and $\theta : 3 \sim 7\,\text{HZ}$. Similarly, the movements *elbow flexion* and *elbow extension* display distinctive patterns at $\theta : 3 \sim 7\,\text{HZ}$ and $\beta : 13 \sim 200\,\text{HZ}$, while the movements *forearm supination* and *forearm pronation* feature unique classifier patterns at $\alpha : 7 \sim 13\,\text{HZ}$. These visualization results demonstrate that the proposed framework is capable of generating distinct classifier patterns for various upper limb

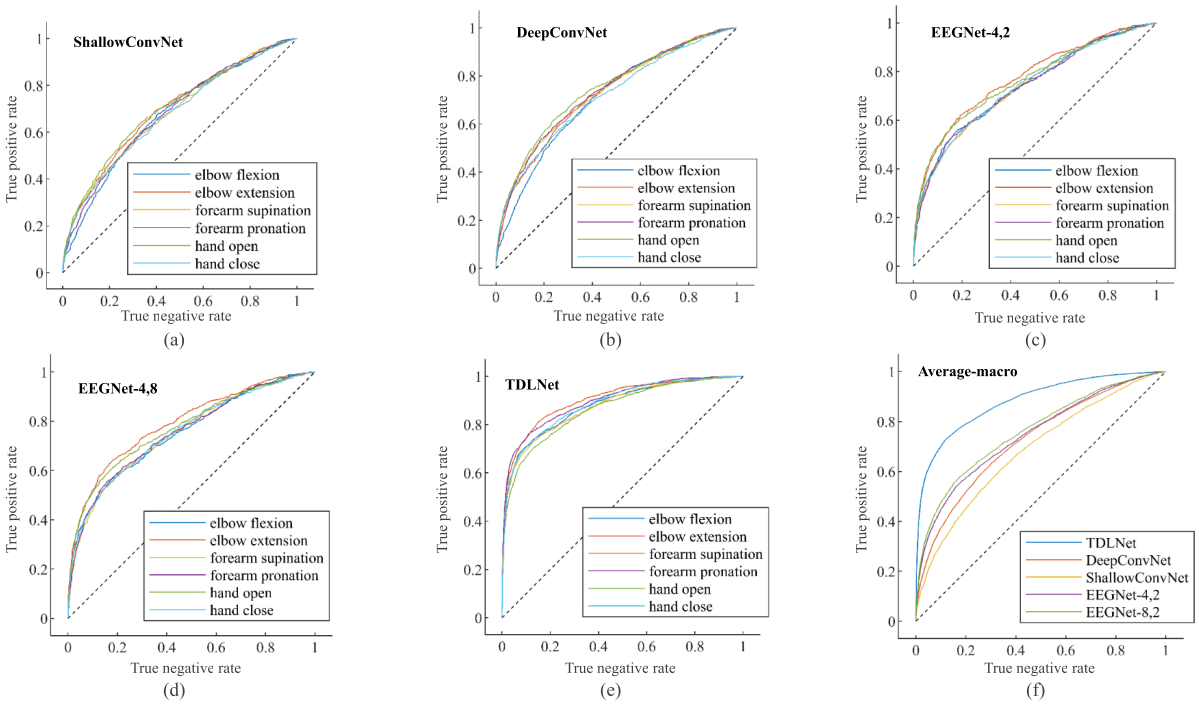

Fig. 4. ROC curves. (a) ShallowConvNet, (b) DeepConvNet, (c) EEGNet-4,2, (d) EEGNet-8,2, (e) TDLNet, (f) Average-macro.

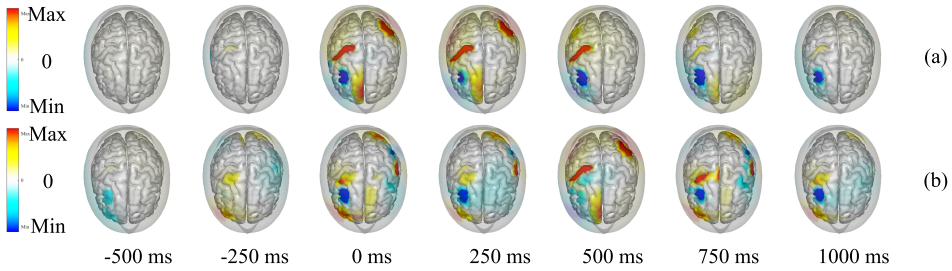

Fig. 5. Visualization of EEG source estimation. (a) ULM6, (b) GRAZ.

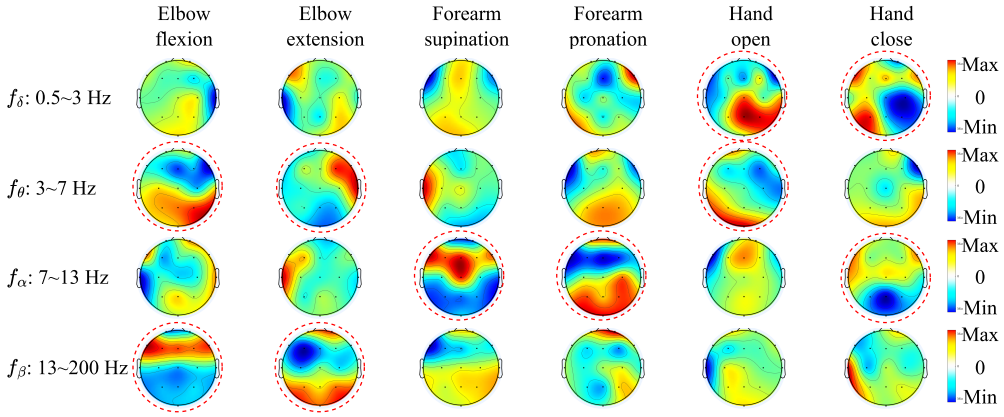

Fig. 6. Feature visualization based on occlusion signal frequency.

motor imagery categories across different frequency bands in EEG signals.

3) Visualization of Feature Before and After RAMM: To illustrate the impact of RAMM in TDLNet, we present Fig. 7, which depicts feature visualization before and after RAMM. Noticeable changes in scalp topographic maps at different EEG channel locations are evident before and after RAMM incorporation. These visualization results highlight that the proposed RAMM can adaptively focus on the most relevant EEG signal channels for a specific task and assign appropriate weights to these channels. This observation aligns with the analysis of brain activation patterns.

IV. Discussion

Unlike traditional non-cross-subject EEG signal classification, which creates personalized models for each user,

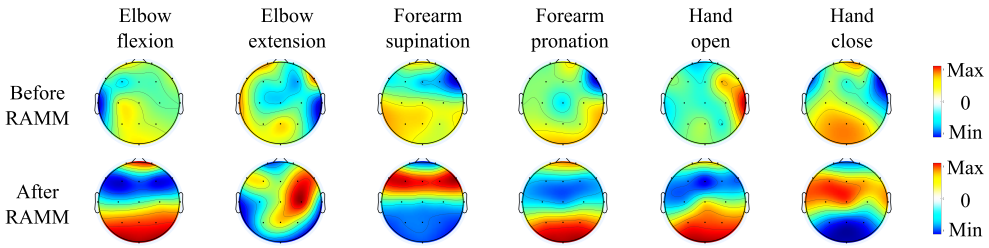

Fig. 7. Visualization of feature before and after RAMM.

cross-subject EEG signal classification promotes the classification of EEG signals among different individuals. The main challenge of cross-subject brain-computer interface research is the differences in brain activity patterns, neural representations, and spatial organization between different individuals, making it difficult to develop universal brain-computer interface algorithms that can be applied to different users. In this study, a novel CNN architecture called TDLNet was proposed to achieve cross-subject classification of six types of motor imagery EEG signals. It mainly consists of three parts. The first part is the TDM module, which is used to process cross-subject electroencephalogram (EEG) signals in groups and then fuse cross-subject channel features through two one-dimensional convolutions, aiming to improve its performance in motor imagery classification by overcoming differences between subjects. The second part is the Inception Module, which uses multiple parallel convolutional branches with different kernel sizes to extract multi-scale temporal information from the input feature map, and the different scales features are connected in the depth direction. The third part is the Residual Attention Mechanism Module, which assigns weights to each EEG signal channel and dynamically focus on the EEG signal channels most relevant to a specific task. Compared with other CNN-based methods (DeepConvNet, ShallowConvNet [38] and EEGNet-4,2, EEGNet-8,2 [39]) and transfer learning methods (MSFBCNN [34], EEGSym [35], MSATNet [36]), TDLNet achieved the best classification results in six categories of upper limb motor imagery EEG signal classification, as shown in TABLE IV and Fig. 4. From TABLE V, it can be observed that, except for ShallowConvNet, other CNN methods combined with TDM achieved accuracy close to that of the transfer learning method, indicating that the TDM module can overcome the differences between subjects by fusing cross-subject subject channel characteristics. The Feature Visualization Algorithm was proposed to qualitatively analyze TDLNet. It can be seen from Fig. 6 that the TDLNet can generate different classifier patterns through signals of different frequencies. The results shown in Fig. 7 can also indicate that RAMM can adaptively pay attention to the EEG signal channels most relevant to a specific task and assigns weights to associated EEG signal channels, resulting in more pronounced classifier patterns on the scale topographical maps.

V. Conclusion

In this article, the Transfer Data Learning Network (TDLNet) is proposed to achieve the cross-subject intention recognition for multiclass upper limb motor imagery. TDLNet is consists of three key components: Transfer Data Module (TDM), Inception Module, and Residual Attention Mechanism Module (RAMM). The TDM processes cross-subject EEG signals in groups and fuses cross-subject channel features using two one-dimensional convolutions. The Inception Module captures multi-scale temporal information from the input feature graph through multiple parallel branches. The RAMM assigns weight to each EEG signal channel, while dynamically focusing on the EEG signal channels most relevant to a specific task. Subsequently, we introduce a feature visualization algorithm based on occlusion signal frequency to qualitatively analyze TDLNet. Ablation studies are conducted to demonstrate the necessity of each module in TDLNet. Comparing TDLNet with CNN-based reference methods (DeepConvNet, ShallowConvNet and EEGNet-4,2, EEGNet-8,2) and transfer learning methods (MSFBCNN [34], EEGSym [35], MSATNet [36]) on the ULM6 and the GRAZ dataset, TDLNet achieves the best classification results. Experimental results confirm that TDLNet successfully achieves cross-subject intention recognition for multiple categories of upper limb motor imagery.

References

[1] U. Chaudhary, N. Birbaumer, and A. Ramos-Murguialday, "Brain–computer interfaces for communication and rehabilitation," *Nature Rev. Neurol.*, vol. 12, no. 9, pp. 513–525, 2016.

[2] K. K. Ang and C. Guan, "Brain–computer interface for neurorehabilitation of upper limb after stroke," *Proc. IEEE*, vol. 103, no. 6, pp. 944–953, Jun. 2015.

[3] X. Gao, Y. Wang, X. Chen, and S. Gao, "Interface, interaction, and intelligence in generalized brain–computer interfaces," *Trends Cognit. Sci.*, vol. 25, no. 8, pp. 671–684, Aug. 2021.

[4] N. Lu, T. Li, X. Ren, and H. Miao, "A deep learning scheme for motor imagery classification based on restricted Boltzmann machines," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 6, pp. 566–576, Jun. 2017.

[5] H. Wang et al., "Discriminative feature extraction via multivariate linear regression for SSVEP-based BCI," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 24, no. 5, pp. 532–541, May 2016.

[6] Z. Shuqfa, A. N. Belkacem, and A. Lakas, "Decoding multi-class motor imagery and motor execution tasks using Riemannian geometry algorithms on large EEG datasets," *Sensors*, vol. 23, no. 11, p. 5051, May 2023.

[7] H. Wang et al., "EEG-based motor imagery recognition framework via multisubject dynamic transfer and iterative self-training," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Feb. 20, 2023, doi: 10.1109/TNNLS.2023.3243339.

[8] R. Chaisaen et al., "Decoding EEG rhythms during action observation, motor imagery, and execution for standing and sitting," *IEEE Sensors J.*, vol. 20, no. 22, pp. 13776–13786, Nov. 2020.

[9] S. Siuly and Y. Li, "Improving the separability of motor imagery EEG signals using a cross correlation-based least square support vector machine for brain–computer interface," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 20, no. 4, pp. 526–538, Jul. 2012.

[10] L. He, D. Hu, M. Wan, Y. Wen, K. M. von Deneen, and M. Zhou, "Common Bayesian network for classification of EEG-based multiclass motor imagery BCI," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 46, no. 6, pp. 843–854, Jun. 2016.

[11] X. Liu, S. Xiong, X. Wang, T. Liang, H. Wang, and X. Liu, "A compact multi-branch 1D convolutional neural network for EEG-based motor imagery classification," *Biomed. Signal Process. Control*, vol. 81, Mar. 2023, Art. no. 104456.

[12] G. C. D. Virgilio, A. J. H. Sossa, J. M. Antelis, and L. E. Falcón, "Spiking neural networks applied to the classification of motor tasks in EEG signals," *Neural Netw.*, vol. 122, pp. 130–143, Feb. 2020.

[13] P. Autthasan et al., "A single-channel consumer-grade EEG device for brain–computer interface: Enhancing detection of SSVEP and its amplitude modulation," *IEEE Sensors J.*, vol. 20, no. 6, pp. 3366–3378, Mar. 2020.

[14] O. Friman, I. Volosyak, and A. Graser, "Multiple channel detection of steady-state visual evoked potentials for brain–computer interfaces," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 4, pp. 742–750, Apr. 2007.

[15] N.-S. Kwak and S.-W. Lee, "Error correction regression framework for enhancing the decoding accuracies of ear-EEG brain–computer interfaces," *IEEE Trans. Cybern.*, vol. 50, no. 8, pp. 3654–3667, Aug. 2020.

[16] Y. Zou, V. Nathan, and R. Jafari, "Automatic identification of artifact-related independent components for artifact removal in EEG recordings," *IEEE J. Biomed. Health Informat.*, vol. 20, no. 1, pp. 73–81, Jan. 2016.

[17] H. Cecotti and A. Graser, "Convolutional neural networks for P300 detection with application to brain–computer interfaces," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 3, pp. 433–445, Mar. 2011.

[18] G. Pfurtscheller and F. H. Lopes da Silva, "Event-related EEG/MEG synchronization and desynchronization: Basic principles," *Clin. Neurophysiol.*, vol. 110, no. 11, pp. 1842–1857, Nov. 1999.

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

[20] H. Huang et al., "An EEG-based brain computer interface for emotion recognition and its application in patients with disorder of consciousness," *IEEE Trans. Affect. Comput.*, vol. 12, no. 4, pp. 832–842, Oct. 2021.

[21] J. Atkinson and D. Campos, "Improving BCI-based emotion recognition by combining EEG feature selection and kernel classifiers," *Expert Syst. Appl.*, vol. 47, pp. 35–41, Apr. 2016.

[22] I. Käthner, S. C. Wriessnegger, G. R. Müller-Putz, A. Kübler, and S. Halder, "Effects of mental workload and fatigue on the P300, alpha and theta band power during operation of an ERP (P300) brain–computer interface," *Biol. Psychol.*, vol. 102, pp. 118–129, Oct. 2014.

[23] F. Fahimi, Z. Zhang, W. B. Goh, T.-S. Lee, K. K. Ang, and C. Guan, "Inter-subject transfer learning with an end-to-end deep convolutional neural network for EEG-based BCI," *J. Neural Eng.*, vol. 16, no. 2, Apr. 2019, Art. no. 026007.

[24] G. D. Flumeri et al., "EEG-based mental workload neurometric to evaluate the impact of different traffic and road conditions in real driving settings," *Frontiers Human Neurosci.*, vol. 12, p. 509, Dec. 2018.

[25] Y. Zhang, C. S. Nam, G. Zhou, J. Jin, X. Wang, and A. Cichocki, "Temporally constrained sparse group spatial patterns for motor imagery BCI," *IEEE Trans. Cybern.*, vol. 49, no. 9, pp. 3322–3332, Sep. 2019.

[26] S. Sakhavi, C. Guan, and S. Yan, "Learning temporal information for brain–computer interface using convolutional neural networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 11, pp. 5619–5629, Nov. 2018.

[27] Y. R. Tabar and U. Halici, "A novel deep learning approach for classification of EEG motor imagery signals," *J. Neural Eng.*, vol. 14, no. 1, Feb. 2017, Art. no. 016003.

[28] O.-Y. Kwon, M.-H. Lee, C. Guan, and S.-W. Lee, "Subject-independent brain–computer interfaces based on deep convolutional neural networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 10, pp. 3839–3852, Oct. 2020.

[29] B. J. Edelman, B. Baxter, and B. He, "EEG source imaging enhances the decoding of complex right-hand motor imagery tasks," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 1, pp. 4–14, Jan. 2016.

[30] J.-H. Jeong, K.-H. Shim, D.-J. Kim, and S.-W. Lee, "Brain-controlled robotic arm system based on multi-directional CNN-BiLSTM network using EEG signals," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 5, pp. 1226–1238, May 2020.

[31] X. Ma, S. Qiu, and H. He, "Time-distributed attention network for EEG-based motor imagery decoding from the same limb," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 30, pp. 496–508, 2022.

[32] B. Yang, J. Ma, W. Qiu, J. Zhang, and X. Wang, "The unilateral upper limb classification from fMRI-weighted EEG signals using convolutional neural network," *Biomed. Signal Process. Control*, vol. 78, Sep. 2022, Art. no. 103855.

[33] R. C. Maswanganyi, C. Tu, P. A. Owolawi, and S. Du, "Multi-class transfer learning and domain selection for cross-subject EEG classification," *Appl. Sci.*, vol. 13, no. 8, p. 5205, Apr. 2023.

[34] H. Wu et al., "A parallel multiscale filter bank convolutional neural networks for motor imagery EEG classification," *Frontiers Neurosci.*, vol. 13, p. 1275, Nov. 2019.

[35] S. Pérez-Velasco, E. Santamaría-Vázquez, V. Martínez-Cagigal, D. Marcos-Martínez, and R. Hornero, "EEGSym: Overcoming inter-subject variability in motor imagery based BCIs with deep learning," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 30, pp. 1766–1775, 2022.

[36] L. Hu, W. Hong, and L. Liu, "MSATNet: Multi-scale adaptive transformer network for motor imagery classification," *Frontiers Neurosci.*, vol. 17, Jun. 2023, Art. no. 1173778.

[37] X. Zhao et al., "Deep CNN model based on serial–parallel structure optimization for four-class motor imagery EEG classification," *Biomed. Signal Process. Control*, vol. 72, Feb. 2022, Art. no. 103338.

[38] R. T. Schirrmeister et al., "Deep learning with convolutional neural networks for EEG decoding and visualization," *Hum. Brain Mapping*, vol. 38, no. 11, pp. 5391–5420, Nov. 2017.

[39] V. J. Lawhern, A. J. Solon, N. R. Waytowich, S. M. Gordon, C. P. Hung, and B. J. Lance, "EEGNet: A compact convolutional neural network for EEG-based brain–computer interfaces," *J. Neural Eng.*, vol. 15, no. 5, Oct. 2018, Art. no. 056013.

[40] P. Ofner, A. Schwarz, J. Pereira, and G. R. Müller-Putz, "Upper limb movements can be decoded from the time-domain of low-frequency EEG," *PLoS ONE*, vol. 12, no. 8, Aug. 2017, Art. no. e0182578.

[41] D. P. Kingma and J. L. Ba, "Adam: A method for stochastic optimization," in *Proc. 3rd Int. Conf. Learn. Represent.*, 2015, pp. 1–15.

[42] R. D. Pascual-Marqui, C. M. Michel, and D. Lehmann, "Low resolution electromagnetic tomography: A new method for localizing electrical activity in the brain," *Int. J. Psychophysiol.*, vol. 18, no. 1, pp. 49–65, Oct. 1994.

[43] F. Tadel, S. Baillet, J. C. Mosher, D. Pantazis, and R. M. Leahy, "Brainstorm: A user-friendly application for MEG/EEG analysis," *Comput. Intell. Neurosci.*, vol. 2011, pp. 1–13, Jan. 2011.