A Cross-Space CNN With Customized Characteristics for Motor Imagery EEG Classification

Ying Hu, Yan Liu, Siqi Zhang, Ting Zhang, Bin Dai, Bo Peng, Hongbo Yang, and Yakang Dai

***Abstract*— The classification of motor imagery-electroencephalogram(MI-EEG)based brain-computer interface(BCI)can be used to decode neurological activities, which has been widely applied in the control of external devices. However, two factors still hinder the improvement of classification accuracy and robustness, especially in multi-class tasks. First, existing algorithms are based on a single space (measuring or source space). They suffer from the holistic low spatial resolution of the measuring space or the locally high spatial resolution information accessed from the source space, failing to provide holistic and high-resolution representations. Second, the subject specificity is not sufficiently characterized, resulting in the loss of personalized intrinsic information. Therefore, we propose a cross-space convolutional neural network (CS-CNN) with customized characteristics for four-class MI-EEG classification. This algorithm uses the modified customized band common spatial patterns (CBCSP) and duplex mean-shift clustering (DMSClustering) to express the specific rhythms and source distribution information in cross-space. At the same time, multi-view features from the time, frequency and space domains are extracted, connecting with CNN to fuse the characteristics from two spaces and classify them. MI-EEG was collected from 20 subjects. Lastly, the classification accuracy of the proposed is 96.05% with real MRI information and 94.79% without MRI in the private dataset. And the results in the BCI competition IV-2a show that CS-CNN outperforms the state-of-the-art algorithms, achieving an accuracy improvement of 1.98%, and a standard deviation reduction of 5.15%.**

***Index Terms*— Brain-computer interface (BCI), motor imagery (MI), cross-space, customized characteristics, convolutional neural network (CNN).**

Manuscript received 27 September 2022; revised 31 January 2023; accepted 23 February 2023. Date of publication 2 March 2023; date of current version 8 March 2023. This work was supported in part by the National Natural Science Foundation of China under Grant 61971413 and Grant 62271481, in part by the Ministry of Science and Technology (MOST) 2030 Brain Project under Grant 2022ZD0208500, in part by the Jiangsu Key Research and Development Plan under Grant BE2021012-5 and Grant BE2022049-2, in part by the Jiangsu International Cooperation Project under Grant BZ2022028, in part by the Jiangsu Province Basic Research Project of Leading Technology under Grant BK20192004, in part by the Soochow Key Basic Research Special Foundation under Grant SJC2022012, and in part by the International Partnership Program of Chinese Academy of Science under Grant 154232KYSB20200016. *(Corresponding authors: Yakang Dai; Yan Liu.)*

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Suzhou Science City Hospital under Application No. IRB202207004RI.

Ying Hu and Bin Dai are with the Division of Life Sciences and Medicine, School of Biomedical Engineering (Suzhou), University of Science and Technology of China, Hefei 230026, China, and also with the Suzhou Institute of Biomedical Engineering and Technology, Chinese Academy of Sciences, Suzhou, Jiangsu 215163, China (e-mail: ying_h_bme@163.com).

Yan Liu is with the Department of Medical Image, Suzhou Institute of Biomedical Engineering and Technology, Chinese Academy of Sciences, Suzhou, Jiangsu 215163, China, and also with Suzhou Guoke Kangcheng Medical Technique Company Ltd., Suzhou 215163, China (e-mail: liuyan@sibet.ac.cn).

Siqi Zhang is with the School of Mechanical and Power Engineering, Harbin University of Science and Technology, Harbin 150080, China.

Ting Zhang is with the School of Automation, Harbin University of Science and Technology, Harbin 150080, China.

Bo Peng and Hongbo Yang are with the Suzhou Institute of Biomedical Engineering and Technology, Chinese Academy of Sciences, Suzhou, Jiangsu 215163, China.

Yakang Dai is with the Suzhou Institute of Biomedical Engineering and Technology, Chinese Academy of Sciences, Suzhou, Jiangsu 215163, China, and also with Suzhou Guoke Kangcheng Medical Technology Development Company Ltd., Suzhou 215163, China (e-mail: daiyk@sibet.ac.cn).

Digital Object Identifier 10.1109/TNSRE.2023.3249831

I. Introduction

BRAIN-COMPUTER interface (BCI) is a technology that enables information exchange between the brain and external devices by constructing pathways that are independent of peripheral nerves and muscle [1]. It has broad application prospects in the fields of entertainment, industrial process, aerospace, and rehabilitation engineering [2] and [4]. At present, non-invasive electroencephalogram (EEG) BCI is widely used for recording brain activity in the field of BCI. Because it can monitor large-scale neuronal activity in the entire brain adjacent to the cranium in a low-cost and risk-free manner [5]. There are four commonly used experimental paradigms for non-invasive BCI: steady-state visual evoked potential (SSVEP), visual P300, error-related potential (ERP), and motor imagery (MI) [6], [7], [8], [9]. In contrast, MI has the outstanding advantages of low damage and signal stability, which has become one of the most promising paradigms [10]. MI-based BCI refers to judging the subject's intention by the activation effect of different brain regions when the subject imagines a specific limb or muscle movement. This process is accompanied by event-related desynchronization (ERD) and synchronization (ERS) of the cerebral cortex based on distinct frequency band mu rhythm (8-12Hz) and beta rhythm (13-30Hz) [11].

Traditional MI-EEG classification algorithms are based on multi-channel EEG in scalp space to identify motor intent, also known as measuring space-based decoding. In recent years, methods for signal classification by extracting the time, frequency, and space domain characteristics of multi-

channel MI-EEG have been developed to a certain extent [12]. Specifically, three-dimensional control of the virtual helicopter was accomplished by directly extracting the difference in the mu rhythm spectrum amplitude of the left and right EEG signals [13]. Despite the rapid development of measuring space-based decoding, various limitations still exist. Although EEG has an extremely high temporal resolution, its low spatial resolution is a fatal problem, which cannot be resolved by expanding the number of electrodes. Since the signals recorded by each EEG electrode are the result of the coupling of multiple intracranial nerve sources, it is not accurate enough to express the detailed spatial features of MI-EEG, which restricts the further improvement of the classification accuracy [14]. Thus, how to enhance the spatial resolution of EEG signals for BCI is a great challenge.

To overcome the mentioned above problems and achieve more accurate detection of the relationship between MI-EEG and motor intentions, it is necessary to develop a neural activity recognition technique. EEG source imaging (ESI) technology provides a solution that maps the signal on the scalp to the source distribution of the cortex to complete the traceability of the EEG, while retaining high temporal resolution, but also improving spatial resolution [15], [16]. A large number of equivalent dipole signals are used to simulate the source signals of the cerebral cortex. What's more, due to the addition of magnetic resonance imaging (MRI) brain anatomical constraints in ESI, source signals also have clearer and more precise physical and physiological interpretation [14]. At present, the method of feature extraction from cortical source signals has been given increasingly more consideration and broadly utilized in BCI motor neural decoding, which is called source space-based decoding [17], [18]. Compared with traditional decoding methods, source space methods have been shown to improve the classification accuracy of MI tasks [19], [20], [21]. Although anatomical techniques and neurophysiological studies provide the theoretical basis for source space-based algorithms, ESI is essentially a model-based neuroimaging technique that has limitations not to be ignored [22]. For one thing, the source signal is calculated rather than real; for another, to avoid the over-fitting problem caused by a large number of dipoles, it is necessary to manually screen out the dipoles highly relevant to MI tasks [18]. Consider the whole, the brain network is an interconnected and coordinated whole. Local brain regions after screening cannot represent the global information of the cerebral cortex, which may lose vital information. Even though the source space-based decoding methods have many advantages and have shown certain superiority over measuring space-based methods, it is a challenge to break through their unreal and local limitations.

The mentioned above facts indicate that the single-space decoding is inadequate for the interpretation of MI-EEG signals, failing to provide holistic and high-resolution representations. In detail, measuring space provides real global information, and source space provides calculated and detailed information. Both of them have their advantages and can complement each other. As far as we know, no research has proposed the fusion of two spaces to decode MI-BCI tasks.

Another problem faced by the classification methods in both measuring and source space is that the classification performance of subjects may differ greatly due to their different brain control and learning ability [23], [24]. It has been pointed out that classification algorithms respond differently to different frequency bands, that is to say, the signals of each subject are specific to the frequency domain [25]. Similarly, it has been proved that the locations of brain activation areas are not the same when subjects perform the MI tasks, which means the signals are specific to the space domain [26]. As a consequence, if the differences between subjects are not considered, the consistent processing of all signals based on physical cognition may lead to problems such as poor robustness, low reliability, and difficulty in further enhancing the accuracy.

To break through the single-space decoding one-sidedness and limitations of subject specificity, we proposed a cross-space convolutional neural network (CS-CNN), in which individual differences were taken into account. Customized characteristics were extracted in the measuring and source space, respectively, and were fed into CNN for cross-space fusion to ultimately achieve four-class MI tasks classification.

This paper innovatively proposed the CS-CNN for decoding MI-EEG four-class tasks, and its contributions can be summarized as the following two points:

- We proposed a novel cross-space decoding, which combined the real global information in measuring space and the simulated detailed information in source space to express the features related to MI tasks more comprehensively and deeply. This had profound implications for improving the accuracy of decoding algorithms.
- We proposed a customized characteristics representation to mine the specificity of MI-EEG in the frequency and space domains, further enhancing the robustness of personalized decoding.

The remainder of this paper is organized as follows. Section II introduces some backgrounds and related works. Section III explains the data sources and the structure of the proposed algorithm. Section IV presents the results of the proposed algorithm and comparative experiments, Section V is the discussion, and Section VI is the relevant conclusion.

II. Related Work

This section first introduces the MI-EEG decoding methods based on measuring space, including machine learning and deep learning. And then, we introduce the source space-based decoding. Finally, we give a comprehensive overview of the proposed algorithm.

A. Decoding Methods Based on Measuring Space

Measuring space-based decoding has been part of mainstream thinking for the past few decades of history, with researchers intervening from both machine learning and deep learning perspectives for in-depth exploration.

In the field of machine learning, some studies have proposed classical methods to classify MI tasks. The performance mainly depends on two important points: distinguishing features and excellent classifiers. Among the numerous features, spatial filters can maximize feature differentiation, such as beamforming, laplacian, and common space pattern

(CSP) [27]. Of particular note, the CSP is the most popular method with obvious advantages [28]. And one versus one (OVO) or one versus rest (OVR) strategies can extend it for multi-task classification. Later, this method has been developed to a certain extent through adaptive CSP, regularized CSP, and L1 norm-solved CSP in enhancing the signal-to-noise ratio and solving the overfitting problem [29], [30], [31]. In particular, the filter bank common spatial patterns (FBCSP) algorithm divides EEG signals into several frequency bands through bandpass filtering, which achieves a great deal [32]. The continuous validation of publicly available datasets has confirmed the undoubted advancement of the FBCSP algorithm and has led to a gradual increase in interest in methods combining filtering and CSP [33], [34], [35]. Existing methods use a priori knowledge to fix the optimal filter range or filter bank, which improves the efficiency of the algorithm. And Xu et al. [36] used a search tree to determine the optimal narrow band to add personalized frequency domain information, providing an idea to solve the subject specificity problem. Support vector machine (SVM) and linear discriminant analysis (LDA) are recognized as the most commonly used classifiers to handle the mentioned above features [37].

With the rapid strides that have been made in deep learning, it can achieve satisfactory classification results without or with simply feature extraction, offering a potentially attractive approach in the field of MI-based BCI. Currently, deep learning such as convolutional neural networks (CNN), recurrent neural networks (RNN), long short-term memory (LSTM), and gated recurrent neural networks (GRNN) have been applied to the MI-EEG classification in measuring space [38], [39], [40], [41]. Amin et al. [43] proposed a lightweight network for extracting EEG dynamic spatial context information and time series by inception-attention module and Bi-LSTM. Luo et al. [41] generated time slices through a clipping strategy to form spatial-frequency-sequence relations, which were incorporated into RNN for classification. Autthasan et al. [42] integrated deep metric learning into an autoencoder to build end-to-end multi-task learning, which provided innovative solutions without pretreatment. Besides, it is worth mentioning that CNN is considered to be one of the most promising networks for solving classification problems in the MI-based BCI field [44]. For example, end-to-end CNN was used to extract temporal and spatial information from EEG signals [45]. Multi-branch CNN was used to classify three-dimensional representations of EEG signals [46]. Although the above methods reduce the burden of feature extraction, poor interpretability and structural lightness dictate that they cannot differentiate subject-specific information and there is little room for model robustness improvement.

B. Decoding Methods Based on Source Space

It was not until the emergence of source space analysis methods based on ESI technology that dominance of measuring space as the dominant means of analysis was broken. However, due to a large number of sources, it would be extremely burdensome to introduce them all into the subsequent calculation. Therefore, it becomes crucial to focus on the effective information from thousands of sources. For example, Fang et al [14] transformed signal classification into an image classification problem by mapping cortical activation at specific frequencies. And returning to signal processing itself, researchers have chosen to label the sensorimotor cortex, which is closely related to MI, as the region of interest (ROI) on a physiological basis. Noirhomme et al. [47] reconstructed the source with a simple head model and manually extracted the dipole of the motor cortex to classify the MI-EEG. Hou et al. [20] created 10 scouts according to gyri markers in the motor cortex to form the ROI and extracted Morlet wavelet features.

As mentioned in part I, numerous works have confirmed the superiority of source space decoding compared to the measuring. However, it achieves high resolution while losing the wholeness of using all channels. At the same time, existing ROI selection ways do not take into account the issue of inter-subject-specific differences in source distribution, and we believe that it is more promising to take full advantage of the spatial activation of source signals to form data-driven ROIs.

To combine the advantages of measuring and source space decoding, also take full account of subject variability, we proposed a CS-CNN. In the measuring space, CSP features after customized EEG rhythms description were extracted. In the source space, customized ROI time series in a highly activated state are formed by data-driven. The global information of the measuring space and the detailed information of the source space were input into CNN simultaneously. In the convolution process, the intrinsic pattern of the signal was further extracted to finish the cross-space fusion, and ultimately achieved high-precision classification of four MI tasks.

III. Methods

This section first describes our private dataset and public dataset. After that, we elaborate proposed algorithm which includes feature extraction of two spaces and a cross-space fusion network. Finally, we present a comparative protocol for evaluating experimental results as a whole.

A. Data Description

We verified our proposed algorithm with a private dataset and BCI competition IV-2a dataset [48]. The experimental paradigm adopted by the self-test dataset is shown in Fig. 2. Other details are explained as follows:

1) Private Dataset: 20 non-disabled volunteers, aged 30±6, participated in the experiment. There are 12 males and 8 females. They are all right-handed and have no BCI experience. All subjects who volunteered for this study received informed written consent. The EEG signals are recorded by a system with 64 active electrodes (BioSemi B.V., Amsterdam, Netherlands) at a sampling frequency of 512 Hz. The electrode locations are shown in Fig. 1. The system replaces the ground electrode with two separate CMS/DRL electrodes. The gel is injected to reduce the impedance of electrodes. In addition, in order to discuss the effect of the presence or absence of real MRI information, we divide all subjects into two groups of 10 individuals each, with subjects' T1-weighted MRIs with 1mm thickness provided by Suzhou Science City Hospital in

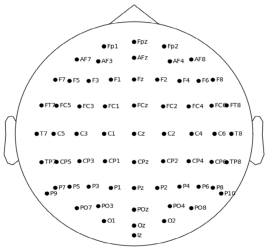

Fig. 1. 64-channel electrode positions.

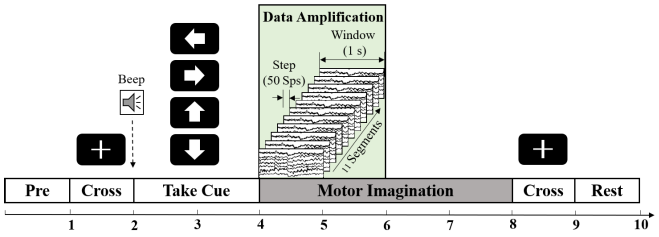

Fig. 2. Experimental paradigm for the private dataset. A sliding window strategy was adopted to expand the data for the first 2 s of the MI task. The Sps represents the samples.

the private dataset 1 and no MRI information in the private dataset 2.

The subjects sit in a comfortable position in front of the computer and are instructed to avoid blinking, eye-rolling, and any form of body movement while executing the MI tasks. Fig. 2 shows the paradigm adopted by the experiment. Each trial lasts 10 s. First, there is 1 s for preparation, followed by a white fixation cross on the black screen and a beep that lasts for 500 ms to prompt the subjects to pay attention to the appearance of the task sign. Next, up/down/left/right arrows appear randomly, instructing subjects to perform the corresponding MI tasks of tongue curl/feet up/left hand in fist/right hand in fist. The cue disappears after 2 s, and at the same time, subjects engage in motor imagination for 4 s. The imagination ends when the fixation cross reappears on the screen, and subjects enter a relaxed resting state waiting for the start of the next trial. Each subject's experiments consist of three blocks, spaced a day apart to ensure that the brain is fully rested. Each block contains 4 sessions, with an interval of 3 minutes to avoid excessive visual fatigue. Each session contains 60 trials, with each of the four task threads appearing 15 times. Each subject performs a total of 720 ($60 \times 4 \times 3$) trials.

2) BCI Competition IV-2a Dataset: It is one of the most commonly used public datasets in the field of MI-BCI and contains data from 22 EEG electrodes(FC3,FC1,FCz,FC2,FC4, C5,C3,C1,Cz,C2,C4,C6,CP3,CP1,CPz,CP2,CP4,P1,Pz and P2) sampled at 250 Hz for 9 subjects. And the data has been preprocessed by bandpass filtering of 0.5∼100 Hz [48]. Consistent with our dataset, it contains four MI tasks (left hand, right hand, tongue, and feet). In this paradigm, subjects perform MI for 3 s. The data contains 2 sessions on different days, designated as training set and test set respectively. Each session contains 288 trials, thus a total of 576 (288×2) trials are performed for each subject.

B. Data Preprocessing

EEGLAB toolbox is used to preprocess signals, including 50 Hz notch filtering, 0.1-32 Hz bandpass filtering, independent component analysis (ICA) for ocular artifacts removal, and baseline correction. The mentioned above operations are carried out in sequence. In addition, data amplification is a crucial part of the training of deep models [49]. The sliding window strategy through time series clipping has been proven to improve MI-EEG classification performance effectively [33]. Considering that subjects may not be able to keep their attention for a long time during the experiment, we select a 2 s period after the beginning of imagination to expand data in the way of setting a sliding window. The step length is 50 sampling points, the window width is 1 s, and the time series is divided into 11/6 segments for the private/public dataset, as shown in Fig. 2. The same preprocessing operation is performed for both datasets. Formally, we define $C \times T$ a single-trail filtered MI-EEG, where C is the number of channels and T is the number of sampled time points. The impact of sliding windows on classification performance will be discussed in section IV. A.

C. Proposed Architecture

An overview of our proposed CS-CNN is illustrated in Fig. 3. This algorithm consists of three modules as a whole: the measuring space module, the source space module, and the cross-space module. The pre-processing operation is considered to be part of the measuring space module which is also shown in Fig. 3. The specific structure of the CNN is drawn in Fig. 4.

1) Measuring Space Module: In measuring space, we propose a customized band common spatial pattern (CBCSP) to first study the inherent customized EEG rhythms by wavelet packet decomposition (WPD), and then followed by the globally spatial representations by CSP. The CBCSP consists of three parts as shown below.

Step 1: MI-EEG sub-band calculation. Inspired by [50], we utilize WPD to decompose signals into several high and low-frequency sub-bands subtly. As displayed in Fig. 2, the preprocessed MI-EEG (0.1-32 Hz) is decomposed by a four-layer wavelet packet, and 16 sub-frequency bands with a width of 2 Hz are acquired. Daubechies wavelet basis is adopted, and the filter length is 4. Sub-band energy coefficient is defined by $\boldsymbol{E}_j$, where j is the node. It's defined by

$$\boldsymbol{E}_j = \sqrt{\sum_k \left|c_j(k)\right|^2} \Big/ \sum_{j=1}^{16} \sqrt{\sum_k \left|c_j(k)\right|^2} \tag{1}$$

where c_j is the wavelet packet coefficient of the j^{th} sub-band, and k is the sampling point.

Step 2: Customized MI-EEG connected-band selection. We select the specific MI-EEG rhythms using the seed-growing method. The connected band consists of several sub-bands dependently. We set $E_m = max(E_j)$ as the seed, and the growth direction alternates between downward (higher frequency) and upward (lower frequency). The neighboring

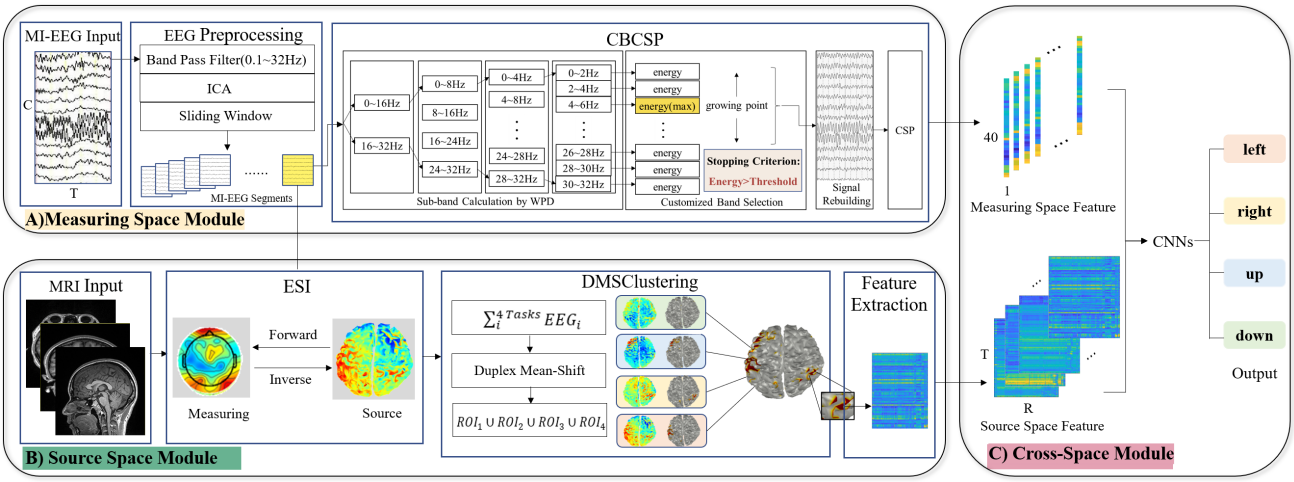

Fig. 3. An overview of proposed CS-CNN. (A) displays preprocessing operation of MI-EEG input and CBCSP in measuring space. (B) exhibits the process of ESI of MRI input and DMSClustering which is responsible for extracting the high-dimensional feature matrix of source space. And the features of measuring space and source space are deeply mined and fused in (C). The structure of CNN is shown in Fig. 4.

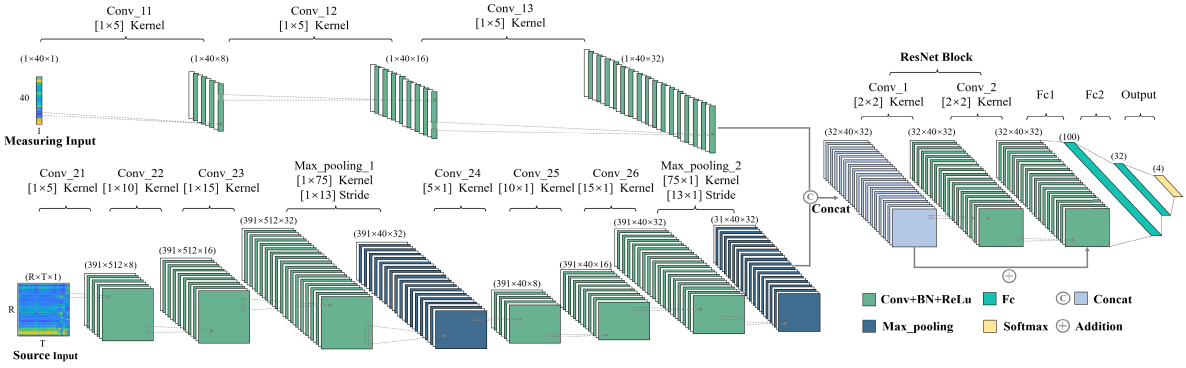

Fig. 4. The illustration of the proposed CNN architecture.

bands are gradually absorbed to form a connected-band set $\boldsymbol{V}$. The growth stops when the total energy is greater than the threshold δ. The growth process is shown in (2).

$$\boldsymbol{V} = \left\{ c_{m+s} \left| \sum_{s=0,1,-1,2,-2,\ldots} E_{\mathrm{m}+s} > \delta \right. \right\} \tag{2}$$

where s is the displacement.

Connected-band in $\boldsymbol{V}$ is reconstructed to form a new MI-EEG with customized rhythmical information. We set δ in (2) as 0.90, which has been proved to be the optimal parameter by our previous study [51].

Step 3: Global spatial representations. We use OVRCSP to maximize the gap between different MI-EEGs for the representation of global spatial information. The basic idea is to construct two kinds of signals by analogy with the traditional CSP and transform the four-class tasks into four binary classification tasks. Define $s, s' \in \{1, 2, 3, 4\}$ as the MI-EEG of four types, where s contains one class of signals and s' contains the remaining three classes of signals. For example, when $s = 2$, $s' = \{1, 3, 4\}$. In this way, the best spatial feature distinguished from the other classes of class 2 can be obtained. Spatial features of four categories are calculated in turn and joined to shape the overall OVRCSP features. We set the optimal number of spatial filter pairs as 5, so the OVRCSP feature dimension of a single-trail MI-EEG is # spatial filter $\times 2 \times$ # category number ($5 \times 2 \times 4$).

2) Source Space Module: In source space, we focus on the customized spatial characteristics in detail and time-related information which are not represented in measuring space by mapping scalp MI-EEG to the cortical source space based on ESI.

Step 1: ESI. It converts MI-EEG into neuronal activity with the high spatial resolution by combining EEG with precise information on head anatomy and source localization algorithms. ESI involves the solution of two crucial problems: the forward problem and the inverse problem. The forward problem is responsible for determining the potential of each scalp electrode generated by known sources in the brain. Firstly, personalized real head models and source models are constructed based on the anatomical information provided by each subject's MRI for the private dataset 1. And the common template in Brainstorm is used for the public dataset and private dataset 2. Secondly, the registration of electrode distribution and head model is carried out manually. Finally, we use the finite element method (FEM) to solve the lead field matrix which is in charge of converting scalp EEG [52]. The solution to the inverse problem is carried out after the forward

problem. The relationship can be expressed as (3), which $\boldsymbol{M}$ represents MI-EEG, $\boldsymbol{L}$ the lead field, $\boldsymbol{N}$ the measuring noise, and C the source signal:

$$\boldsymbol{M} = \boldsymbol{LC} + \boldsymbol{N} \tag{3}$$

The equivalent current dipole model shows good applicability in solving the inverse problem by equivalenting local neurons in the brain to a current dipole with several parameters (position and pole phase). Thus, we choose the weighted minimum norm estimation (WMNE) to calculate the distribution of cortical current dipoles [53]. 7344 (5001) dipoles are obtained for private (public) datasets. Finally, we complete the mapping of MI-EEG from scalp measuring space to source space, which improves the spatial resolution of signals and provides more detailed spatial information for subsequent classification.

Step2: Customized ROI extraction. We propose a duplex mean-shift clustering (DMSClustering) to screen out the customized ROI, in which dipoles are highly relevant to MI tasks. Before DMSClustering, the four types of source signals are superimposed separately for sub-ROIs clustering of different MI-tasks. And the final ROI is the union of the four sub-ROIs.

DMSClustering is a data-driven method to cluster dipoles by considering both activation intensity and distribution location, ensuring that the ROI is sufficiently active and concentrated clusters. We establish an abstract spherical window with the core c and radius r by simultaneous constraints of distance factor S_q and intensity factor I_t, as shown in (4), (5), and (6).

$$S_q = \frac{\|x_i - x_c\|_2 + \|y_i - y_c\|_2 + \|z_i - z_c\|_2}{\sum_{i=1}^{N} (\|x_i - x_c\|_2 + \|y_i - y_c\|_2 + \|z_i - z_c\|_2)} \tag{4}$$

$$I_t = \|A_i - A_c\|_2 \Big/ \sum_{i=1}^{N} (\|A_i - A_c\|_2) \tag{5}$$

$$\boldsymbol{S_k} = \left\{(x, y, z, A) \,\middle|\, S_q(x, y, z) + I_t(A) < r\right\} \tag{6}$$

where (x, y, z) represents the position coordinates of dipoles, A represents activation intensity, and N the total number of dipoles. The radius r is set to 0.1.

The core drifts towards the mean value of all dipoles' intensities in the window until it converges. Each drift is determined by an average displacement vector $\boldsymbol{M}$, whose direction is always from the highly activated state to the hypo-activated state, which is expressed in (7).

$$\boldsymbol{M} = \frac{1}{n} \sum_{x_i \in \boldsymbol{S_k}} (A(x_i) - A(x_c)) \tag{7}$$

where n the number of dipoles collected in $\boldsymbol{S_k}$.

We take the dipole with the highest activation intensity as the initial core c, and the clustering process ends when the number of visited dipoles exceeds the percentile of N.

Step 3: Spatiotemporal characterization. The ROI time sequences are clipped by the same sliding window strategy in measuring space to obtain several high-dimensional spatiotemporal feature matrices. Without any algorithmic processing, they provide the most discriminating time and space scales of original details. For example, the source space feature dimension of a single-trail signal of subject 1 in the private dataset is # ROI dipoles × # window length ×1 (316×512×1).

3) Cross-Space Module: The module takes charge of the hidden information extraction and fusion of the mentioned above feature matrices from measuring space and source space by a multi-layer CNN. The network structure is shown in Fig. 4.

In the measuring branch, we adopt a shallow network structure that contains three convolutional layers to further learn the implicit global spatial feature information. We select a 1D convolution kernel (1×5) to match the input size (1×40), and its length is equivalent to that of the CSP spatial filter. Following the learning rules from low complexity to high complexity, the number of three convolution kernels is set to 8, 16, and 32 respectively. There is no feature map compression during the whole process, so the size of the feature matrix in the measuring space after convolutional is still (1 × 40).

In the source space branch, spatially separable convolution and variable receptive field strategies are used to better match the network input characteristics. Inspired by the MobileNet, we split the $n \times n$ convolution kernel into $1 \times n$ and $n \times 1$ for extracting the detailed temporal and spatial information provided by cortical source input respectively [54]. In addition, we set up three convolution layers composed of kernels with different sizes in both time and space directions to prevent repeated learning in local areas from redundancy, n = 5, 10, 15. It is foreseeable that with the increase of the receptive field, the more information obtained, the better global features obtained. The varied receptive fields can learn rich and diverse features at different levels. The number of convolution kernels per three is the same as in the measuring space. Two maximum pooling layers are set respectively after three consecutive convolution layers of time and space to reduce the amounts of data and parameters and prevent over-fitting. The sizes of two pooling kernels are (1 × 75) and (75 × 1), and strides are 13. The additional purpose is to reduce the length of the input matrix in the time series to 40 through dimensionality reduction, which can match the length of the measuring feature and facilitate subsequent feature fusion. Activation locations represented by higher intensity values in the spatiotemporal feature map are important classification criteria, so max pooling provides better performance than average. In general, the network structure which is responsible for deep information mining in source space is composed of six convolution layers and two maximum pooling layers. Finally, taking subject 1 as an example, the dimension of the source space feature matrix after the multi-layer convolution operation is (38 × 40).

The concatenate layer is adopted to fusion two feature matrices of the measuring space and source space. To realize the re-mining of the deep-level information of the splicing matrix, a residual network (ResNet) module is added after the fusion matrix. It can directly transfer the shallow features to the deeper layers, ensuring that the original information is not lost while achieving rapid feedback and feature fusion. The size of the convolution kernel in this module is (2 × 2). And the last two fully connected layers are used to integrate the aforementioned features and map the learned feature representation to the sample marker space. Ultimately, the four-class classification is achieved by relying on the softmax layer

to map the outputs of the four neurons $[y^1, y^2, y^3, y^4]$ into the interval of (0,1). Softmax's output represents the relative probabilities between different categories, it can be defined as follows:

$$S_i = e^{y^i} \Big/ \sum_{i=1}^{4} e^{y^i} \tag{8}$$

Another point to note is that we do not want the original feature matrix to incur dimensionality reduction losses except for the pooling operation, so the stride sizes of all convolutional layers are set to 1 in both width and height.

Batch normalization is used to prevent the over-fitting phenomenon in network training, avoid the problem of gradient disappearance, and speed up network convergence. The features after convolution are normalized so that the input of each layer can keep the same distribution. The total training samples are divided into smaller batches. Parameters are updated after completing a batch of sample learning. To find the best balance between memory efficiency and memory capacity, we set the batch size to 16. The rectified linear units (ReLu) activation function is chosen because of the advantage of sparsity. It can be expressed as:

$$\text{Re } Lu(x) = \begin{cases} x & if\ x > 0 \\ 0 & if\ x \leq 0 \end{cases} \tag{9}$$

The cross-entropy loss function is used to measure the difference between the prediction effects. We use the Adam optimization algorithm as the optimizer and minimize the loss function with an initial network learning rate of 1×10^{-5}. The maximum epoch is set as 128, and the learning rate of every 9 epochs is attenuated by a factor of 0.9 to ensure accelerated convergence in the early stage and stable performance in the later stage.

D. Experimental Evaluation

For the matter of four classification tasks, accuracy, Kappa value, and confusion matrix were used to evaluate classification results. The merging of all trials for each subject ensured a five-fold cross-training to eliminate randomisation of results. Wilcoxon rank sum test analysis was used to determine whether there was a significant difference. To evaluate the proposed CS-CNN on three datasets (private dataset 1 and 2, BCI competition IV-2a), we set up the following comparative experiments.

1) Experiment I: Parameter Comparison: To explore the influence of the sliding window strategy on the experimental results of data amplification, we set three different sliding window steps for comparative experiments based on the proposed CS-CNN. Since the EEG signal length was fixed for each subject in a single trial (2 s), longer sliding window steps meant fewer sliding windows, in other words, fewer amplifications of the MI-EEG. The sliding window step was in the unit of the sampling point. We used the accuracy to judge the merits of the four classification results on three datasets.

2) Experiment II: Customized Characteristics Verification: One of the highlights of this paper is paying attention to the intrinsic difference in EEG rhythms and source distribution of subjects. Therefore, to further verify the superiority of the customized characteristics in this paper, which can also be regarded as the robustness of the algorithm, the following three groups of ablation experiments were set up.

Group1: CSP with fixed bandpass filtering instead of CBCSP in the measuring space module and cortical motor area instead of customized ROI in the source space module.

Group2: customized ROI is instead of motor region in the source space module.

Group3: CSP with fixed bandpass filtering instead of CBCSP in the measuring space module.

3) Experiment III: Space-Based Classification Performance: To verify the superiority of cross-space compared to single-space decoding, the measuring and source space modules are ablated to compare classification performance of the three datasets. Single-space based decoding consists of two groups, with only one module (measuring or source) used for feature extraction, and the classification network consists of the corresponding branch directly connected to the fully connected and output layers, discarding the concatenate connection and ResNet module. The cross-space based decoding refers to the proposed CS-CNN algorithm.

4) Experiment IV: Algorithm Performance Verification: To verify the superiority of CS-CNN without the impact of data quality, we collected the state-of-the-art studies using BCI Competition IV-2a data for four classifications in recent years as the baseline.

FBCSP-RNN: Luo et al. [41] applied FBCSP to extract features and then used a sliding cropping technique to generate spatial-frequency-sequences for input into RNN.

NSL-EEGNet: Raza et al. [55] combined neural structured learning (NSL) and EEGNet to standardize neural network training by using relational information in data.

MBCNN: Altuwaijri and Muhammad, [39] classified EEG signals without any processing by adopting a multi-branch CNN (MBCNN)model with different convolution kernels.

3DCNN: Zhao et al. [46] generated a 3D representation by keeping the MI-EEG in a sequence of 2D arrays of the spatial distribution of sampled electrodes. For the 3D representation, a multi-branch 3D CNN was designed.

Inception-CNN: Zhang et al. [56] amplified the data by adding noise to be fed in an improved CNN containing inception and ResNet modules.

CNN-LSTM: Amin et al. [43] proposed a CNN framework based on the inception-attention mechanism to extract spatial context information and dynamic features and connected bi-LSTM which was responsible for time series information.

SHNN: Liu et al [58] proposed to learn different time window CSP features and perform sparse representation for classification by a SincNet-hybrid neural network (SHNN).

IV. Results

This section presents the results and statistical analysis of experiments I, II, III, and IV to validate the effectiveness.

A. Experiment I

Fig. 5 shows the influence of different sliding window steps on classification accuracy in three datasets. Solid diamonds represent the accuracy of all subjects in each dataset.

TABLE I
CUSTOMIZED EEG RHYTHMS AND SOURCE DISTRIBUTION IN THREE DATASETS

Data	Attribute	1	2	3	4	5	6	7	8	9	10	General
Private Dataset 1	Frequency band/Hz	8-24	8-26	8-26	8-28	8-20	6-26	8-26	6-24	8-32	8-20	8-30
	Dipoles	160	270	297	340	284	180	325	331	229	391	260
Private Dataset 2	Frequency band/Hz	6-30	8-28	6-26	8-26	10-28	6-26	8-24	6-20	6-32	6-30	8-30
	Dipoles	107	106	164	185	115	205	123	146	136	194	125
BCI Competition IV-2a	Frequency band/Hz	6-18	6-24	6-26	8-16	6-22	6-20	6-20	6-20	6-26	-	8-30
	Dipoles	221	266	168	266	223	187	147	283	171	-	125

TABLE II
CLASSIFICATION PERFORMANCE(ACCURACY ± SD IN % AND KAPPA ±SD) OFCUSTOMIZED CHARACTERISTICS ABLATION EXPERIMENT

Dataset	Method	Cross-space	Customized rhythms	Customized ROI	Accuracy (%)	Kappa	SDA (%)	ARE (%)
Private Dataset 1	Group1	√			93.19±1.06	0.91±0.02	3.59	12.43
	Group2	√	√		95.13±0.87	0.94±0.02	1.75	5.74
	Group3	√		√	94.40±1.55	0.92±0.02	2.39	7.31
	CS-CNN	√	√	√	**96.05±0.62**	**0.95±0.02**	**1.33**	**3.79**
Private Dataset 2	Group1	√			89.82±2.34	0.88±0.02	2.69	7.97
	Group2	√	√		92.76±0.88	0.91±0.02	1.76	5.14
	Group3	√		√	92.73±1.34	0.91±0.02	1.67	4.58
	CS-CNN	√	√	√	**94.79±1.18**	**0.93±0.02**	**1.08**	**3.90**
BCI competition IV-2a	Group1	√			84.74±2.43	0.82±0.02	7.84	22.05
	Group2	√	√		86.13±1.78	0.84±0.02	6.61	18.74
	Group3	√		√	87.25±1.88	0.84±0.02	4.70	13.09
	CS-CNN	√	√	√	**90.80±2.37**	**0.88±0.02**	**3.41**	**9.59**

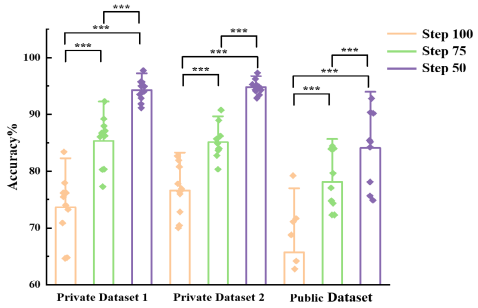

Fig. 5. Accuracy of different sliding window steps of three datasets. The solid diamond represents the classification accuracy for each subject.

We observed that as the sliding window step decreased from 100 sampling points to 50, the classification accuracy showed an upward trend gradually. And when marking the step as 50, the CS-CNN achieved the best performance for all datasets ($p<0.001$). It is proved that data amplification is a powerful approach to improving the final classification accuracy. As a result, we choose the sliding window step size to be 50 and used it in the subsequent experiments.

B. Experiment II

Table I shows the customized EEG rhythms and source distribution. They are displayed visually by the specific band range screened by CBCSP and the number of dipoles obtained by DMSClustering. In addition, Table I includes the general frequency bands and the number of dipoles in the motor cortex for comparison. Note that, the selected bandwidths were narrower than the general one on the whole. And lower frequency bands (6Hz) were screened in nearly half of the results, which were considered highly relevant to the MI tasks. The total number of activated dipoles per subject fluctuated within a certain range.

Table II illustrates the classification performance of the customized characteristics ablation experiment in the three benchmark datasets, using customized and general characteristics displayed in Table I. In this regard, the standard deviation of inter-subject accuracy (SDA) and the inter-subject accuracy range error (ARE) were calculated to assess the robustness of the CS-CNN. It can be found that the proposed CS-CNN algorithm achieved the highest classification accuracy and Kappa value, as well as the lowest SDA and ARE value in all three datasets compared to Group 1, Group 2, and Group 3 ($p<0.05$). It also demonstrated that both the proposed customized EEG rhythm and ROI approach worked.

C. Experiment III

Table III shows the classification performance of three datasets based on space-based decoding. Consistent with the findings of existing studies, we validated that source space decoding outperformed measuring space. In addition, the cross-space decoding reached the best performance in all datasets as we expected, with the two private datasets showing significant differences ($p<0.05$).

TABLE III
CLASSIFICATION PERFORMANCE(ACCURACY ± SD IN % AND KAPPA ±SD) OFSPACE-BASED DECODING. BOLD DENOTES THE BEST NUMERICALVALUES

Dataset	Measuring space-based		Source space-based		Cross space-based (CS-CNN)	
	Accuracy (%)	Kappa	Accuracy (%)	Kappa	Accuracy (%)	Kappa
Private Dataset 1	90.96±0.92	0.88±0.02	94.25±0.94	0.92±0.02	**96.05±0.62**	**0.95±0.02**
Private Dataset 2	88.49±0.86	0.85±0.01	90.79±1.22	0.87±0.01	**94.79±1.18**	**0.93±0.02**
Public Dataset	80.21±2.64	0.76±0.02	89.55±1.82	0.86±0.02	**90.80±2.37**	**0.88±0.02**

TABLE IV
CLASSIFICATION PERFORMANCE (ACCURACY IN % AND KAPPA) OF PROPOSED COMPAREDTO THE STATE-OF-THE-ART METHODS IN BCI COMPETITION IV-2A DATASET. BOLD DENOTES THE BEST NUMERICAL VALUES

	FBCSP-RNN[41]	NSL-EEGNet[55]	MBCNN[39]	3D CNN[46]	Inception-CNN[56]	CNN-LSTM[43]	SHNN[58]	CS-CNN
1	84.82	82.29	82.58	77.40	89.61	89.23	82.76	**91.72**
2	65.32	51.39	70.01	60.14	80.01	72.53	68.97	**88.48**
3	83.54	85.07	93.79	82.93	96.17	**97.23**	79.31	91.72
4	67.67	67.01	82.60	72.29	81.26	76.28	65.52	**88.95**
5	64.00	58.33	77.81	75.83	83.76	82.48	58.62	**88.31**
6	70.87	56.25	64.79	68.99	81.20	69.15	48.28	**89.12**
7	84.96	83.33	88.02	76.04	94.75	**94.76**	86.21	89.53
8	71.95	73.96	86.91	76.86	**98.28**	86.14	89.66	91.78
9	68.90	78.47	83.83	84.67	90.50	86.10	89.87	**93.75**
Mean	73.56	70.68	81.15	75.02	88.39	82.84	74.26	**90.37**
SDA	8.04	12.84	9.04	7.34	7.06	9.62	14.77	**1.91**
ARE	20.96	33.68	29.00	24.53	18.27	25.61	41.59	**5.44**

D. Experiment IV

Table IV shows the overall performance comparison between the CS-CNN and the state-of-the-art algorithms for four-class tasks based on the BCI competition IV-2a. The training and test sets fixed by the data provider were used in this experiment. It can be found that CS-CNN outperformed all other algorithms with average accuracy (90.37%), Kappa value (0.88), SDA (1.91%), and ARE (5.44%). In detail, the proposed improved average accuracy by at least 1.98% and reduced the SDA by at least 5.15% compared to the most advanced algorithm.

Fig. 6 shows violin plots of classification accuracy for three datasets, with the horizontal line being the mean and the solid diamond showing the distribution for each subject. The kernel density plot of the external parcel has a higher probability of being distributed around a value the larger the area of the graph in a region. Due to the strict control of data collection conditions, the distributions of the two private datasets are similar and the accuracy of group 1 is slightly better than that of group 2. The public dataset is more dispersed and slightly less well classified.

Fig. 7 displays the confusion matrix of the private and public datasets. The matrices are plotted based on the average of all subjects' classification results in each dataset. It can be seen that the corresponding overall accuracies of the left hand, right hand, tongue, and feet were 96.68%, 96.45%,96.60%, and 95.28% in the private dataset 1. And they were 93.20%, 97.07%, 95.64% and 93.48% in the private dataset 2. And they were 91.18%, 92.42%, 87.45% and 92.30% in public dataset. By integrating the results of three matrices, we can find that in the four-class tasks, the mean classification accuracy of the tongue and feet MI was slightly lower than the other two,

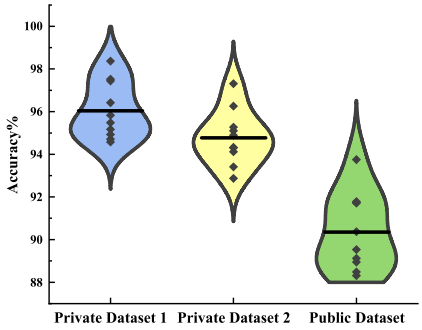

Fig. 6. Violin plot of accuracy for the three datasets, horizontal line represents the mean, solid diamond is the classification accuracy for each subject.

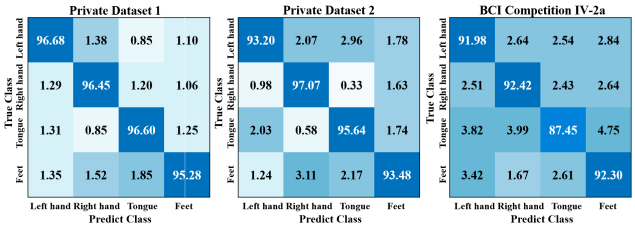

Fig. 7. Confusion matrices of four-class MI-EEG classification.

and the likelihood of their mutual confusion was higher as well.

V. DISCUSSION

A. Effectiveness of Sliding Window Strategy

The most direct and operative way to counteract the over-fitting problem of deep learning is by amplifying the data volume. In particular, the restricted amount of data in the

public dataset limits the proposal of deeper and finer network architecture, so most of the relevant research has adopted data amplification strategies [43], [46], [56]. As shown in Fig. 4, the sliding window strategy used for data amplification has a great influence on the classification results. It reveals that utilizing the sliding window to extract the sample during the period of comparatively concentrated attention of the subjects (2 s) is an efficient way to portray the time-variant information, which is conducive to subsequent network learning with more discriminative features.

B. Analysis of Customized Characteristics

One of the contributions of this paper is to characterize the frequency and spatial specificity simultaneously. Numerous studies have proved that the MI-EEG is closely related to the ERD/ERS from the cortex of the brain, especially the motor area. According to Table I, we can find that although each subject shares common characteristics, like the current physiological findings, they have their own diversity as well. This diversity is reflected in narrower and lower frequency bands, or a more restricted or wider source distribution than in the cortical motor area. The impact of these customized characteristics on the classification results is also verified in Table II. Compared with consistent processing using general characteristics, customized EEG rhythms and ROI optimize feature representation by fully mining the implicit frequency and spatial activation information inherent to each subject, allowing the maximization of valuable information from the MI-EEG. Focusing on the public dataset with variable data quality, an important reason for the impressive performance of CS-CNN is the outstanding contribution of customized ROI. It has been clearly indicated that the brain activity of people who are not good at BCI is not only linked to the motor area, but also has complex connections to multiple brain regions related to cognition and emotion [57]. It is reasonable to speculate that this enhancement is due to the data-driven customized source distribution-based approach that pays good attention to activation information outside the motion area that is easily overlooked. Overall, the customized characteristics can learn and improve the discrimination of MI-EEG among different classes, which has great significance for achieving a personalized and robust BCI system.

C. Analysis of Cross-Space Performance

In general, the measuring space provides real and holistic signal representation, and the source space can interpret the neural activity information in a detailed view. There are certain complementarities between them. Cross-space decoding achieves the optimization of comprehensive multi-view feature representation by fusing the advantages of both, thus solving the limitation of single-space decoding. According to Table III, the classification performance based on different spaces is in the order of cross-space, source space, and measuring space, and the advantage of cross-space decoding is fully demonstrated in the private dataset. However, the significance analysis conducted for each subject showed that the CS-CNN classification results for subjects 2, 4, 5 and 6 in the public dataset were closer to the source space-decoding, ($p>0.05$). The reason is that the traceability analysis occupies the most significant role in improving the classification results, and the source space module's interpretation of brain states has maximized the valuable information in the MI-EEG, making the fusion less important. This can be further illustrated in conjunction with the excellent performance of the customized ROI in Table II based on the above subjects' data. However, it is undeniable that in most cases cross-space decoding still shows a strong competitive edge. To sum up, CS-CNN provides a powerful new approach for MI four-tasks decoding, which has absolute power to surpass the widely used measuring space decoding and has considerable potential to perform better than the emerging source space decoding.

D. Analysis of Overall Algorithm Structure

Our proposed CS-CNN achieves global high-resolution feature extraction through cross-space fusion to improve the accuracy of the algorithm and enhances robustness through customized feature characterization to address subject specificity issues. According to Table IV, although the proposed algorithm did not achieve the best classification accuracy for the case of good data quality (subjects 3, 7, 8), it did significantly improve the results for the case of poor BCI ability (subjects 2, 5, 6). By balancing the performance of different subjects' data, CS-CNN obtained the strongest robustness while accommodating high accuracy. Overall, based on the comparison results with a variety of other representative algorithms (including FBCSP-RNN, improved EEGNet, innovative CNN, hybrid networks, etc.), it is confirmed that our proposed can achieve optimal classification performance under several evaluation metrics. This proves that the classical structure of feature representation connected to deep networks can still show strong competitiveness in the field of MI-based BCI decoding in the present day when various new hybrid networks are gradually emerging.

According to Fig. 5, the classification results for private dataset 1 were excellent, but the need for real MRI information makes it more suitable for the field of BCI-based intelligent rehabilitation. Surprisingly, the private dataset 2 without MRI information also showed quite competitive results. This gives the CS-CNN a wider and more lifelike application scenario.

E. Future Direction

Although CS-CNN has shown exciting results, there are still some immature areas that need further study in the future. First, more in line with the MI-EEG characteristics and more intelligent data amplification strategies that are needed for the training of high-performance classifiers. Second, more complete strategies are required for cross-space feature fusion, such as considering the attention mechanism to offset the negative effects of some poor-quality data or features, to ensure the stable performance of the fusion structure. Third, a more lightweight network structure is desired for large-scale use and training of BCI in practical task scenarios.

VI. Conclusion

In this study, a cross-space fusion algorithm with customized characteristics is proposed to resolve four classification tasks of MI-EEG. In this algorithm, we perform multi-view feature expressions in the time domain, frequency domain, and space domain (detailed and global) to mine more comprehensive and richer MI-EEG information. On top of a fine-grained interpretation of neural activity in source space, CS-CNN adds real information in measuring space, which provides a new avenue for MI-EEG classification. Finally, the CS-CNN achieved an accuracy of 96.05% and Kappa value of 0.95 in the group with MRI information, and 94.79% accuracy, Kappa 0.93 in the group without MRI in the private dataset four classification task. In BCI competition IV-2a, the average accuracy rate is 90.37%, and the Kappa value is 0.88. Compared with the most advanced algorithm, the accuracy rate has increased by 1.98% and the SDA value has decreased by 5.15%, which shows that our proposed is a high-precision and robust model for multi-task MI-EEG classification.

References

[1] J. Wolpaw, N. Birbaumer, D. McFarland, G. Pfurtscheller, and T. Vaughan, "Brain-computer interfaces for communication and control," *Clin. Neurophys.*, vol. 113, no. 6, pp. 767–791, Aug. 2002.

[2] A. K. Singh, Y.-K. Wang, J.-T. King, and C.-T. Lin, "Extended interaction with a BCI video game changes resting-state brain activity," *IEEE Trans. Cogn. Develop. Syst.*, vol. 12, no. 4, pp. 809–823, Dec. 2020.

[3] M. Xu, J. Han, Y. Wang, T.-P. Jung, and D. Ming, "Implementing over 100 command codes for a high-speed hybrid brain-computer interface using concurrent P300 and SSVEP features," *IEEE Trans. Biomed. Eng.*, vol. 67, no. 11, pp. 3073–3082, Nov. 2020.

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

[5] S. Waldert, "Invasive vs. non-invasive neuronal signals for brain-machine interfaces: Will one prevail?" *Frontiers Neurosci.*, vol. 10, pp. 265–269, Jun. 2016.

[6] Y. Chen, C. Yang, X. Ye, X. Chen, Y. Wang, and X. Gao, "Implementing a calibration-free SSVEP-based BCI system with 160 targets," *J. Neural Eng.*, vol. 18, no. 4, Jun. 2021, Art. no. 046094.

[7] W. Wei, S. Qiu, X. Ma, D. Li, B. Wang, and H. He, "Reducing calibration efforts in RSVP tasks with multi-source adversarial domain adaptation," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 11, pp. 2344–2355, Nov. 2020.

[8] A. Ravi, J. Lu, S. Pearce, and N. Jiang, "Enhanced system robustness of asynchronous BCI in augmented reality using steady-state motion visual evoked potential," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 30, pp. 85–95, 2022.

[9] J. Jin et al., "A novel classification framework using the graph representations of electroencephalogram for motor imagery based brain-computer interface," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 30, pp. 20–29, 2022.

[10] R. Abiri, S. Borhani, E. W. Sellers, Y. Jiang, and X. Zhao, "A comprehensive review of EEG-based brain–computer interface paradigms," *J. Neural Eng.*, vol. 16, no. 1, Feb. 2019, Art. no. 011001.

[11] H. Yuan and B. He, "Brain–computer interfaces using sensorimotor rhythms: Current state and future perspectives," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 5, pp. 1425–1435, May 2014.

[12] F. Lotte, M. Congedo, A. Lécuyer, F. Lamarche, and B. Arnaldi, "A review of classification algorithms for EEG-based brain–computer interfaces," *J. Neural Eng.*, vol. 4, no. 2, pp. R1–R13, Jun. 2007.

[13] A. J. Doud, J. P. Lucas, M. T. Pisansky, and B. He, "Continuous three-dimensional control of a virtual helicopter using a motor imagery based brain-computer interface," *PLoS ONE*, vol. 6, no. 10, Oct. 2011, Art. no. e26322.

[14] T. Fang et al., "Decoding motor imagery tasks using ESI and hybrid feature CNN," *J. Neural Eng.*, vol. 19, no. 1, Feb. 2022, Art. no. 35078158.

[15] H. Yuan, T. Liu, R. Szarkowski, C. Rios, J. Ashe, and B. He, "Negative covariation between task-related responses in alpha/beta-band activity and BOLD in human sensorimotor cortex: An EEG and fMRI study of motor imagery and movements," *NeuroImage*, vol. 49, no. 3, pp. 2596–2606, Feb. 2010.

[16] B. Kamousi, A. N. Amini, and B. He, "Classification of motor imagery by means of cortical current density estimation and von Neumann entropy," *J. Neural Eng.*, vol. 4, no. 2, pp. 17–25, Jun. 2007.

[17] F. Lotte, A. Lecuyer, and B. Arnaldi, "FuRIA: An inverse solution based feature extraction algorithm using fuzzy set theory for brain–computer interfaces," *IEEE Trans. Signal Process.*, vol. 57, no. 8, pp. 3253–3263, Aug. 2009.

[18] C. Li, H. Guan, Z. Huang, W. Chen, J. Li, and S. Zhang, "Improving movement-related cortical potential detection at the EEG source domain," in *Proc. 10th Int. IEEE/EMBS Conf. Neural Eng. (NER)*, May 2021, pp. 214–217.

[19] 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.

[20] Y. Hou, L. Zhou, S. Jia, and X. Lun, "A novel approach of decoding EEG four-class motor imagery tasks via scout ESI and CNN," *J. Neural Eng.*, vol. 17, no. 1, Feb. 2020, Art. no. 016048.

[21] A. Seeland, M. M. Krell, S. Straube, and E. A. Kirchner, "Empirical comparison of distributed source localization methods for single-trial detection of movement preparation," *Frontiers Hum. Neurosci.*, vol. 12, p. 340, Sep. 2018.

[22] V. S. Handiru, A. P. Vinod, and C. Guan, "EEG source imaging of movement decoding: The state of the art and future directions," *IEEE Syst., Man, Cybern. Mag.*, vol. 4, no. 2, pp. 14–23, Apr. 2018.

[23] M. Ahn and S. C. Jun, "Performance variation in motor imagery brain–computer interface: A brief review," *J. Neurosci. Methods*, vol. 243, pp. 103–110, Mar. 2015.

[24] P. J. García-Laencina, G. Rodríguez-Bermudez, and J. Roca-Dorda, "Exploring dimensionality reduction of EEG features in motor imagery task classification," *Expert Syst. Appl.*, vol. 41, no. 11, pp. 5285–5295, Sep. 2014.

[25] C. R. Hema, M. P. Paulraj, S. Yaacob, A. H. Adom, and R. Nagarajan, "An analysis of the effect of EEG frequency bands on the classification of motor imagery signals," *Int. J. Biomed. Soft Comput. Hum. Sci., Off. J. Biomed. Fuzzy Syst. Assoc.*, vol. 16, no. 1, pp. 121–126, Jan. 2010.

[26] M. Li, C. Zhang, S. Jia, and Y. Sun, "Classification of motor imagery tasks in source domain," in *Proc. IEEE Int. Conf. Mechatronics Autom. (ICMA)*, Aug. 2018, pp. 83–88.

[27] M. Grosse-Wentrup, C. Liefhold, K. Gramann, and M. Buss, "Beamforming in noninvasive brain–computer interfaces," *IEEE Trans. Biomed. Eng.*, vol. 56, no. 4, pp. 1209–1219, Apr. 2009.

[28] H. Ramoser, J. Müller-Gerking, and G. Pfurtscheller, "Optimal spatial filtering of single trial EEG during imagined hand movement," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 8, no. 4, pp. 441–446, Dec. 2000.

[29] X. Song and S. C. Yoon, "Improving brain-computer interface classification using adaptive common spatial patterns," *Comput. Biol. Med.*, vol. 61, pp. 150–160, Jun. 2015.

[30] F. Lotte and C. Guan, "Regularizing common spatial patterns to improve BCI designs: Unified theory and new algorithms," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 2, pp. 355–362, Feb. 2011.

[31] P. Li, P. Xu, R. Zhang, L. Guo, and D. Yao, "L1 norm based common spatial patterns decomposition for scalp EEG BCI," *Biomed. Eng. OnLine*, vol. 12, no. 1, pp. 77–87, Aug. 2013.

[32] K. K. Ang, Z. Y. Chin, H. Zhang, and C. Guan, "Filter bank common spatial pattern (FBCSP) in brain–computer interface," in *Proc. IEEE Int. Joint Conf. Neural Netw.*, Jun. 2008, pp. 90–97.

[33] P. Gaur, H. Gupta, A. Chowdhury, K. McCreadie, R. B. Pachori, and H. Wang, "A sliding window common spatial pattern for enhancing motor imagery classification in EEG-BCI," *IEEE Trans. Instrum. Meas.*, vol. 70, 2021, Art. no. 4002709.

[34] S. Kumar, R. Sharma, and A. Sharma, "OPTICAL+: A frequency-based deep learning scheme for recognizing brain wave signals," *PeerJ Comput. Sci.*, vol. 61, pp. 150–160, Feb. 2021.

[35] K. K. Ang, Y. Zheng, C. Wang, C. Guan, and H. Zhang, "Filter bank common spatial pattern algorithm on BCI competition IV datasets 2a and 2b," *Frontiers Neurosci.*, vol. 6, p. 39, Jan. 2012.

[36] S. Xu, L. Zhu, W. Kong, Y. Peng, H. Hu, and J. Cao, "A novel classification method for EEG-based motor imagery with narrow band spatial filters and deep convolutional neural network," *Cogn. Neurodynamics*, vol. 16, no. 2, pp. 379–389, Apr. 2022.

[37] F. Lotte et al., "A review of classification algorithms for EEG-based brain–computer interfaces: A 10 year update," *J. Neural Eng.*, vol. 15, no. 3, Jun. 2018, Art. no. 031005.

[38] D.-Y. Lee, J.-H. Jeong, B.-H. Lee, and S.-W. Lee, "Motor imagery classification using inter-task transfer learning via a channel-wise variational autoencoder-based convolutional neural network," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 30, pp. 226–237, 2022.

[39] G. A. Altuwaijri and G. Muhammad, "A multibranch of convolutional neural network models for electroencephalogram-based motor imagery classification," *Biosensors*, vol. 12, no. 1, p. 22, Jan. 2022.

[40] Z. Tang, C. Li, and S. Sun, "Single-trial EEG classification of motor imagery using deep convolutional neural networks," *Optik*, vol. 130, pp. 11–18, Feb. 2017.

[41] T.-J. Luo, C.-L. Zhou, and F. Chao, "Exploring spatial-frequency-sequential relationships for motor imagery classification with recurrent neural network," *BMC Bioinf.*, vol. 19, no. 1, p. 344, Sep. 2018.

[42] P. Autthasan et al., "MIN2Net: End-to-end multi-task learning for subject-independent motor imagery EEG classification," *IEEE Trans. Biomed. Eng.*, vol. 69, no. 6, pp. 2105–2118, Jun. 2022.

[43] S. U. Amin, H. Altaheri, G. Muhammad, W. Abdul, and M. Alsulaiman, "Attention-inception and long- short-term memory-based electroencephalography classification for motor imagery tasks in rehabilitation," *IEEE Trans. Ind. Informat.*, vol. 18, no. 8, pp. 5412–5421, Aug. 2022.

[44] H. Altaheri et al., "Deep learning techniques for classification of electroencephalogram (EEG) motor imagery (MI) signals: A review," *Neural Comput. Appl.*, pp. 1–42, Aug. 2021.

[45] E. Lashgari, J. Ott, A. Connelly, P. Baldi, and U. Maoz, "An end-to-end CNN with attentional mechanism applied to raw EEG in a BCI classification task," *J. Neural Eng.*, vol. 18, no. 4, Aug. 2021, Art. no. 0460e3.

[46] X. Zhao, H. Zhang, G. Zhu, F. You, S. Kuang, and L. Sun, "A multi-branch 3D convolutional neural network for EEG-based motor imagery classification," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 10, pp. 2164–2177, Oct. 2019.

[47] Q. Noirhomme, R. I. Kitney, and B. Macq, "Single-trial EEG source reconstruction for brain–computer interface," *IEEE Trans. Biomed. Eng.*, vol. 55, no. 5, pp. 1592–1601, May 2008.

[48] C. Brunner, R. Leeb, G. R. Müller-Putz, A. Schlögl, and G. Pfurtscheller, "BCI competition 2008—Graz data set A," pp. 1–6. [Online]. Available: http://www.bbci.de/competition/iv/desc_2a.pdf

[49] 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.

[50] N. Robinson, A. P. Vinod, K. K. Ang, K. P. Tee, and C. T. Guan, "EEG-based classification of fast and slow hand movements using wavelet-CSP algorithm," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 8, pp. 2123–2132, Aug. 2013.

[51] Y. Hu et al., "Multi-task imagery EEG classification based on adaptive time-frequency common spatial patterns combined with CNN," *J. Biomed. Eng.*, vol. 39, no. 6, pp. 1065–1073, Dec. 2022.

[52] Y. Zhang, L. Ding, W. van Drongelen, K. Hecox, D. M. Frim, and B. He, "A cortical potential imaging study from simultaneous extra- and intracranial electrical recordings by means of the finite element method," *NeuroImage*, vol. 31, no. 4, pp. 1513–1524, Jul. 2006.

[53] Q. Wu, W. Yan, X. Shen, Y. Li, and G. Xu, "EEG source reconstruction based on the boundary-element method and weighted minimum norm approaches," *IEEE Trans. Magn.*, vol. 39, no. 3, pp. 1547–1550, May 2003.

[54] A. G. Howard et al., "MobileNets: Efficient convolutional neural networks for mobile vision applications," 2017, *arXiv:1704.04861*.

[55] H. Raza, A. Chowdhury, S. Bhattacharyya, and S. Samothrakis, "Single-trial EEG classification with EEGNet and neural structured learning for improving BCI performance," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2020, pp. 1–8.

[56] C. Zhang, Y.-K. Kim, and A. Eskandarian, "EEG-inception: An accurate and robust end-to-end neural network for EEG-based motor imagery classification," *J. Neural Eng.*, vol. 18, no. 4, Aug. 2021, Art. no. 046014.

[57] K. Kasahara, C. S. DaSalla, M. Honda, and T. Hanakawa, "Basal ganglia-cortical connectivity underlies self-regulation of brain oscillations in humans," *Commun. Biol.*, vol. 5, no. 1, p. 712, Jul. 2022.

[58] C. Liu et al., "SincNet-based hybrid neural network for motor imagery EEG decoding," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 30, pp. 540–549, 2022.