This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIM.2021.3067943, IEEE Transactions on Instrumentation and Measurement

Performance Enhancement of P300 Detection by Multi-Scale-CNN

Hongtao Wang*, *Member, IEEE*, Zian Pei, Linfeng Xu, Tao Xu, Anastasios Bezerianos, *Senior Member, IEEE*, Yu Sun*, *Senior Member, IEEE*, and Junhua Li*, *Senior Member, IEEE*

***Abstract*—P300-based spelling system is one of the most popular brain-computer interface applications. Its success largely depends on performance, including the information transmission rate (ITR) and detection rate (i.e., accuracy). To achieve good performance, we proposed a multi-scale convolutional neural network (MS-CNN) model, which consists of seven layers. First, an upfront dataset was used to train the MS-CNN, aiming to obtain a subject-unspecific model (universal model) for P300 detection. Second, this universal model was adapted by a portion of data derived from a subject to update the model to obtain a subject-specific model by incorporating a transfer learning technique. We applied the proposed model in the BCI Controlled Robot Contest at the 2019 World Robot Conference, and our performance was the best among the teams in the contest. In the contest, ten healthy young subjects were randomly assigned by the contest committee to assess the model. Our model achieved the best P300 detection performance (higher accuracy with less repetition time). The ITR for the subject-unspecific case was 13.49 bits/min while the ITR for the subject-specific case was 12.13 bits/min when the repetitions were fewer than six. It is believed that our method may pave a promising path for taking a further step toward efficient implementation of P300-based spelling system.**

***Index Terms*—Electroencephalogram (EEG), Subject-Unspecific, Subject-Specific, Multi-Scale Convolutional Neural Network (MS-CNN), Event-Related Potential (ERP)**

I. Introduction

BRAIN-computer interfaces (BCI) provide an alternative method for controlling external devices or communicating between human beings and the outside world by translating brain activities into commands or information [1]. To date, different BCI paradigms have been developed, and numerous BCI-based applications have been proposed. For example, a BCI system can translate brain activities into commands to control a cursor [2]. For disabled people, BCI can be used to provide assistance and facilitate their movements. A brain-driven wheelchair has been tested for this purpose [3], [4]. Among the BCI paradigms, event-related potential (ERP)-based BCI is one of the widely used paradigms due to its high reliability. In particular, P300 [5], a decision-making-related positive waveform occurring approximately 300 ms after the onset of an external stimulus (visual, auditory, or tactile stimulus), has been repeatedly adopted in the ERP-based BCI system [6]. This type of BCI has been utilized for TV controlling [7], virtual keyboards [8], and word typing [9]. However, it is not easy to accurately detect P300 from EEG signals due to the low signal-to-noise ratio (SNR), especially for the case of a single trial (i.e., no repetition) [10]. Moreover, the P300 peak is influenced by users' age [11] and the surrounding environment [12]. To mitigate these effects and enhance the SNR, a few trials (repetition) are usually averaged. This strategy was frequently used in P300 detection studies. However, it leads to low efficiency in the information transmission due to a long time caused by the repetition.

Other than the repetition, detection accuracy is the other critical factor for a successful and efficient P300-based BCI. The detection accuracy is determined by both feature extraction and classification. Different feature extraction methods have been proposed in P300-based BCIs. Farwell and Donchin proposed using correlation coefficients between EEG time series and P300 templates as features and achieved a detection accuracy of 80% [13]. To further improve the detection accuracy, Liu et al. utilized band power and stepwise discriminant analysis (SWDA) to calculate P300 amplitude as the features and achieved an improved accuracy in the custom stimulus paradigm [14]. Nonetheless, the algorithm was slow in calculating frequency band power for a long segment. Furthermore, discriminative canonical pattern matching (DCPM) has been successfully applied to ERP identification and has achieved good performance on five datasets [15]. Other methods, such as independent component analysis [16], sensor selection, and channel selection methods, were also developed to improve feature extraction for P300-based BCI [17]–[22].

The methods used in the P300-based BCI system are not only the traditional methods, such as support vector machine (SVM) (e.g., P300 speller [23]), but also the deep learning method. Deep learning is successful in diverse applications, such as image retrieval, speech recognition, and biomedical

This work was supported by the Special Projects in Key Fields Supported by the Technology Development Project of Guangdong Province under Grant (2020ZDZX3018), the Special Fund for Science and Technology of Guangdong Province under Grant (2020182), the Wuyi University and Hong Kong & Macao joint Research and Development Project under Grant (2019WGALH16), the Science Foundation for Young Teachers of Wuyi University under Grant (2018td01), the Jiangmen Brain-like Computation and Hybrid Intelligence R&D Center under Grant ([2018]359, [2019]26), the National Natural Science Foundation of China (61806149, 81801785), Guangdong Basic and Applied Basic Research Foundation (2020A1515010991). (*Hongtao Wang and Zian Pei contributed equally to this work.*), (** indicates corresponding authors: Hongtao Wang, Yu Sun, Junhua Li*)

H. Wang, Z. Pei, L. Xu and T. Xu are with the Faculty of Intelligent Manufacturing, Wuyi University, Jiangmen, 529020, China. (email: nushongtaowang@qq.com).

A. Bezerianos is with the Department of Medical Physics, University of Patras Greece, 26500 Patras, Greece.

Y. Sun is with the Key Laboratory for Biomedical Engineering of Ministry of Education of China, (email: yusun@zju.edu.cn).

J. Li is with the Faculty of Intelligent Manufacturing, Wuyi University, Jiangmen, 529020, China, also with the School of Computer Science and Electronic Engineering, University of Essex, Colchester, CO4 3SQ, UK (email: juhalee.bcmi@gmail.com)

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIM.2021.3067943, IEEE Transactions on Instrumentation and Measurement

TABLE I: Data illustration of the final round in the BCI Controlled Robot Contest.

Group	Training method	Training data	Subject	Testing data
Subject-unspecific	Precontest datasets for universal model	Each: 5 characters, 50 sessions	Six	Each: 5 characters, 2 sessions
Subject-specific	Precontest datasets for universal model, subject-specific data for transfer learning	Each: 5 characters, 1 session	Four	Each: 5 characters, 2 sessions

signal classification [24]–[28]. It is noteworthy that convolutional neural networks (CNNs) have gained substantial interest due to their strong spatial feature extraction ability [29]–[32], which might be promising algorithms for P300 detection [33], [34]. Cecotti first used CNN to realize the classification of P300 and achieved a high recognition rate (90%) with 10 repetitions [35]. When the number of repetitions was 15, the character recognition accuracy was 98%. However, its information transmission rate (ITR) is low because 15 repetitions are required to recognize one character. The best P300-based BCI ensures good performance (high detection accuracy) while maintaining low repetition to obtain a high information transmission rate.

To achieve the above goal, we propose a multi-scale CNN (MS-CNN) model to detect P300 with low repetition. We set up multi-scale kernels because multiple kernels have shown good performance in other applications, such as driving fatigue detection [36], schizophrenia identification [37] and epileptic EEG classification [38]. In addition, there has been an increasing trend in the use of multi-scale convolution in recent years [39]–[43]. These multiple convolution kernels in the proposed model can extract features at different scales and at different time points to capture more complete information contained in the EEG, compared to previous studies [35]. Therefore, this might counteract the negative effect of the reduction in repetition to make a good balance between detection accuracy and the time needed for the detection. We applied the proposed model in the BCI Controlled Robot Contest at the 2019 World Robot Conference (held in Beijing, Aug. 20 − 25, 2019) and received the champion of the P300-based BCI competition[1].

The main contributions of this study are highlighted as follows:

1) The idea of a multi-scale convolution kernel is applied to a deep learning model to improve the information transmission rate (ITR) of P300.
2) Fine-tuning is adopted in this study to transfer the subject-unspecific model to the subject-specific model.
3) Data augmentation was utilized to relieve the imbalance problem between categories.

The remainder of the paper is organized as follows. The experimental paradigm and data collection are introduced first in section 2. This is followed by the details of the methodology in section 3. Subsequently, the results are presented in section 4. Finally, conclusions are drawn in section 5.

[1] http://www.worldrobotconference.com/html/jiqirendasai/chengji/2019/2019/0925/970.html

II. Experimental Protocol

A. Stimuli

The program committee of the BCI Controlled Robot Contest of the 2019 World Robot Conference determined the experimental protocol and provided the data used for model assessment. The P300 paradigm is shown in Fig. 1, which is similar to that used for collecting the International BCI competition III Data sets II [44]. Briefly, a stimulus interface consisting of a matrix of 6×6 characters was used to induce P300 potentials. Rows and columns in the matrix were successively and randomly flashed for 175 ms. Two out of twelve flashes contained the target character (i.e., the target appeared in a particular row and particular column once, respectively). The responses of infrequent target stimuli are different from those of nontarget stimuli, as the infrequent target stimuli induce P300 [45]. A large label denotes a character label and a small label denotes the column or row level (i.e., if the target character is 'B', the large label corresponds to 'B' and the label of the second column and the third row is set to 1 while other labels are set to 0).

B. Data Collection

A Neusen W device (Neuracle Co., Ltd, China) and a cap with 64 active Ag-AgCl electrodes were used to record EEG signals from the scalp. All electrodes were referenced to Cpz and sampled at a rate of 250 Hz. The impedance was kept below 10 kΩ for all electrodes. The data of 59 electrodes were provided in the contest. To keep the same number of electrodes as that of the precontest datasets for performing transfer learning, 57 electrodes (excluding AF7 and AF8) were selected for further processing. More details about transfer learning and the precontest datasets are introduced in a later section.

C. Information Transmission Rate

To quantitatively assess the effectiveness of the algorithms developed by different teams, the information transmission rate (ITR) was employed in the BCI Controlled Robot Contest, which can be calculated using the following equation:

$$ITR = \frac{60}{T}\left[\log_2 Q + P\log_2 P + (1-P)\log_2\left(\frac{1-P}{Q-1}\right)\right], \quad (1)$$

where Q denotes the number of targets, which equals 36 characters in this work. P is the recognition accuracy of character. T indicates the time it takes for character recognition, which is directly influenced by the number of repetitions. Hence, recognition accuracy and the number of repetitions are two crucial parameters. The ITR unit is bits/min. Of note,

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIM.2021.3067943, IEEE Transactions on Instrumentation and Measurement

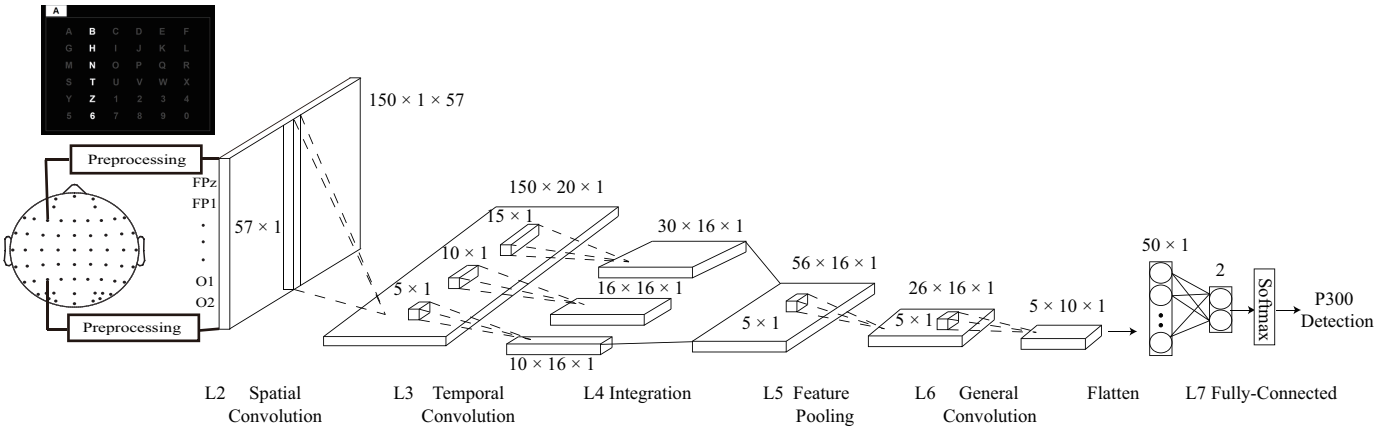

Fig. 1: The architecture of the multi-scale convolutional neural network (MS-CNN). The leaky ReLU activation function and dropout strategy were employed in convolutional layers. Three convolutional kernel scales were used in the MS-CNN.

during the competition, each team was requested to provide the recognition result (target character) within six repetitions. Otherwise, it was considered a failure for the trial.

III. Methodology and Materials

A. Participants

In the competition, ten healthy students (male / female = six / four) were randomly assigned for real-time assessment, six of which were assigned to the subject-unspecific model group, while the remaining four participants were assigned to the subject-specific model group. All participants reported normal or corrected-to-normal vision. In addition, the precontest datasets include two groups. Each group contained 50 sessions, and each session contained five characters.

B. Subject-Unspecific Algorithm

1) Signal Preprocessing: Six subjects participated in the subject-unspecific group. For each repetition, EEG data of five characters were acquired. Then, the obtained EEG data were preprocessed in both temporal-frequency domains. To capture the comprehensive process of P300 occurrence, a segment of 0-600 ms was extracted when the stimulus occurred, resulting in a matrix of 150 (sampling points) × 57 (electrodes). Then the segment was processed with common average reference, detrended, and bandpass at 0.1-20 Hz. These preprocessing steps improved the EEG SNR.

2) Feature Extraction: Given that the amplitude of the P300 signal is small (i.e., at the μv level), it is easily hidden by interference and noise. The average across trials (repetitions) can effectively enhance the signal-to-noise ratio so that the P300 component becomes more visible. Let x be the recorded EEG signal, which could be expressed as the sum of the original signal (s) and the noise (n).

$$x_i(t) = s_i(t) + n_i(t), \tag{2}$$

after averaging across N repetitions, the following equation can be obtained:

$$x_i(t) = \frac{1}{N}\sum_{i=1}^{N} s_i(t) + \frac{1}{N}\sum_{i=1}^{N} n_i(t), \tag{3}$$

assume that the averaged noise is 0, the variance is δ^2, and the noise in different segments is not relevant. The variance after signal averaging is:

$$\begin{aligned}
\delta_e^2 &= E\left[\frac{1}{N}\sum_{i=1}^{N} n_i(t)\right]^2 \\
&= \frac{1}{N^2}\left\{E\left[\sum_{i=1}^{N} n_i(t)^2\right] + 2E\left[\sum_{i=1}^{N}\sum_{j=1}^{N} n_i(t)n_j(t)\right]\right\} \\
&= \frac{1}{N^2}E\left[\sum_{i=1}^{N} n_i(t)^2\right] \\
&= \frac{\delta^2}{N},
\end{aligned} \tag{4}$$

it can be seen that the signal variance after the average was $1/N$ of the original, and the signal-to-noise ratio is significantly improved. In this study, we aimed to achieve a good performance under the fewer repetitions (i.e., smaller N). Therefore, we explored this parameter of N by setting different values to determine its effect on the P300 detection accuracy.

3) Data Augmentation: Data augmentation is an effective approach to improve the identification rate. A more suitable data augmentation method for P300 is proposed based on the previous research on one-dimensional signals [46], [47]. In the experiment, each row and each column were flashed once, and two of them contained the target character. The numbers of target P300 (T-P300) and nontarget P300 (Non-T-P300) were unbalanced (i.e., 1000 (2 × 500) for T-P300 and 5000 (10 × 500) for Non-T-P300). To eliminate the imbalance in the sample number between T-P300 and Non-T-P300, we increased the samples of T-P300 by reinforcing the T-P300

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIM.2021.3067943, IEEE Transactions on Instrumentation and Measurement

TABLE II: The classification accuracy and ITR obtained by the subject-unspecific and subject-specific model.

Modle	Sub.	Repetition(N)											
		1		2		3		4		5		6	
		Acc.	ITR	Acc.	ITR	Acc.	ITR	Acc.	ITR	Acc.	ITR	Acc.	ITR
Subject-unspecific	1	30	11.32	30	7.22	60	16.31	90	25.13	90	20.77	100	21.84
	2	30	11.32	0	0	10	0.64	0	0	20	1.71	30	2.95
	3	20	5.59	70	28.45	60	16.31	80	20.53	70	13.64	70	11.62
	4	0	0	10	0.87	20	2.62	10	0.51	20	1.71	20	1.46
	5	20	5.59	30	7.22	50	12.19	90	25.13	90	20.77	90	17.70
	6	10	1.37	30	7.22	60	16.31	50	9.63	50	7.96	40	4.74
	Avg.	18.33	5.86	28.33	8.50	43.33	10.73	53.33	13.49	56.67	11.09	58.33	10.05
Subject-specific	1	60	34.82	70	28.45	70	20.88	70	16.5	90	20.77	100	21.84
	2	10	1.37	20	3.56	20	2.62	20	2.07	30	3.46	20	1.46
	3	20	5.59	40	11.6	40	8.52	70	16.52	70	13.64	60	9.07
	4	10	1.37	20	3.56	40	8.52	40	6.73	60	10.65	70	11.62
	Avg.	25	10.79	37.5	11.79	42.5	10.14	50	10.45	62.5	12.13	62.5	11

samples. Specifically, the samples obtained by averaging over N (N=2,3,4,5,6) original samples were treated as new samples. In this way, the sample numbers of T-P300 and Non-T-P300 are equal, and the total sample number is 10,000 (i.e., 5,000 for each class).

4) MS-CNN Network: To capture spatial features and temporal features under different scales for P300 detection, an MS-CNN model was proposed (see Fig. 1). The model included seven layers, labeled L1 $\sim$ L7.

L1: Input layer. This layer was used for loading the EEG signal (150$\times$ 1$\times$ 57).

L2: Spatial convolution layer. It consists of a convolutional kernel with a size of 57, which equals the number of signal electrodes. The channel of convolution kernels is 20. This processing method includes weighted superposition averaging and common spatial filtering. It can effectively improve the signal-to-noise ratio of the signal while removing the redundant space information further. The calculation process is as follows (5).

$$x_j^2 = f(\sum_{i \in M_j} I_i \times k_{ij}^2 + b_j^2), \tag{5}$$

where x_j^2 denotes the jth feature map of L2. f is the activation function, using the rectified linear unit (ReLU). I denotes the input data. k is the convolution kernel matrix, and b_j^2 is the additive bias. M_j represents a selection of input maps. Each output map is given an additive bias b; however for a particular output map, the input maps is convolved with distinct kernels. That is, if output map j sum over input map i, then the kernels applied to map i are different for output map j.

L3: Temporal convolution layer. To capture more distinguishing features, three parallel convolutional layers are arranged in the temporal convolution layer. More specifically, the channels of convolutional kernels for three parallel convolutional layers are 16, while the sizes of each kernel are [(5,1), (10,1), (15,1)]. This strategy can help extract diversified temporal information and make these rich features more effective. The calculation process of the temporal convolution layer can be found in formula (6)-(8).

$$x_j^{3,1} = f(\sum_{i \in M_j} x_i^2 \times k_{ij}^{3,1} + b_j^{3,1}), \tag{6}$$

$$x_j^{3,2} = f(\sum_{i \in M_j} x_i^2 \times k_{ij}^{3,2} + b_j^{3,2}), \tag{7}$$

$$x_j^{3,3} = f(\sum_{i \in M_j} x_i^2 \times k_{ij}^{3,3} + b_j^{3,3}), \tag{8}$$

where $x_j^{3,1}$, $x_j^{3,2}$ and $x_j^{3,3}$ represent output maps of different convolution kernels [(5,1), (10,1), (15,1)] in L3. The channel of convolution kernels of each kind is 16.

L4: Integration layer. These feature maps extracted by three temporal filters in L3 are combined. The aim of this layer is to integrate the extracted features.

L5: Feature pooling layer. By the pooling operation, the features obtained in L4 were screened, and the dominant features are selected. The pooling filter size used in this study is (2,1). It helps to reduce computational complexity and prevent overfitting in the context of a small number of training samples.

L6: General convolution layer. This standard convolutional layer includes convolution kernels with a size of (5,1), and the channel of convolution kernels is 10. To extract more abstract, deeper and beneficial features for classification, a convolution filtering operation is carried out on the features obtained by the L5 layer. The calculation process is as follows.

$$x_j^6 = f(\sum_{i \in M_j} x_i^5 \times k_{ij}^6 + b_j^6), \tag{9}$$

where x^5 represents the output of L5. The channel of convolution kernels is 10.

L7: Fully connected layer. Vector x of 1×50 has been reshaped from x^6, and the output value $h_{w,b}(x)$ of the neuron is (10):

$$h_{w,b}(x) = f(w^T x + b), \tag{10}$$

where w^T denotes the weight vector for the fully connected layer. In the decision step, the output of each row and column is determined by the softmax function as a probability. Each row and each column flick once in a round and only two of

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIM.2021.3067943, IEEE Transactions on Instrumentation and Measurement

them contain P300. In other words, P300 only appears when the row and column containing the target character are flicking. As P300 should be unique in row flicking or column flicking, we took the strategy of maximizing probability. The details are shown in the following equations (11)-(12).

$$r = argmax P_r(m_r), m_r \in \{1, 2, \cdots, 6\}, \tag{11}$$

$$c = argmax P_c(m_c), m_c \in \{1, 2, \cdots, 6\}, \tag{12}$$

where r and c denote row and column. P_r and P_c represent the probabilities of P300 appearing in row r and column c. m_r and m_c represent the indices of rows and columns, respectively. P300 appearance is determined by seeking the row and column with maximum probability. Then, the target character can be identified as it should be located in the intersection of the identified row and column.

In this study, the cross-entropy loss function is used to measure the classification error of the network. The regularization method is used in the L2 layer to reduce the risk of overfitting, and the regularization coefficient is set to 0.04. Training weight values with gradient descent, the base learning rate is 0.01, the decay rate is 0.9995, and the maximum number of iterations is 30,000.

C. Subject-specific Algorithm

The subject-specific algorithm is similar to the subject-unspecific algorithm. Both of them are based on the MS-CNN model. The subject-unspecific algorithm builds a universal subject-unspecific model trained by using the precontest datasets. The subject-specific algorithm adapts the universal model by a transfer learning technique using a portion of data collected from that subject. Transfer learning can effectively improve the diagnostic accuracy of the model and has been well applied for experimental bearing vibration signals [48]. In the BCI Controlled Robot Contest, there were four subjects in the subject-specific performance evaluation.

1) Model Adaption: A large quantity of data is required to train a deep learning model with excellent performance. It is not always feasible in practice. When there is a common property existing across datasets, the deep learning model trained in a dataset can be transferred to classify another dataset. If a portion of the data is available to retrain the model, the classification performance can be largely improved. This is the principle of transfer learning [49]. The practice is to train a deep learning model by large EEG data what are related to the transferred classification question. Then, the model is adjusted to meet a new challenge. Fine-tuning is usually used to adjust the parameters of a deep learning model.

The model adaption in this study is fine-tuning, which is based on the universal MS-CNN model. The model structure is retained, and fine-tuning is achieved using subject-specific data. In particular, the subject-specific model is initialized with parameters of the subject-unspecific mode, except for the output layer. The parameters of the output layer were initialized by random values. The output layer is trained with subject-specific data, and the purpose of fine-tuning is achieved through a backpropagation algorithm. We run the backpropagation algorithm for 30,000 iterations, which optimizes the network parameters using adaptive moment estimation. By fine-tuning, the powerful generalization ability of deep neural networks can help avoid complex model design and time-consuming training.

IV. Results and Discussion

A. Model Comparisons

To evaluate the performance of the MS-CNN model, we compared the results to the five representative models (i.e., DCPM, SVM, CNN-FE, CNN-RE and CNN-RM). DCPM is a kind of ERP classification algorithm with high robustness and strong generalization. It is suitable for the condition of a small sample size and can effectively reduce the training time. SVM is a classical machine learning model. CNN-FE, CNN-RE and CNN-RM were proposed in [35] and are the first successful deep learning models for P300 detection.

- CNN-FE (Convolutional Neural Network with Full Electrode): The network is composed of four layers (equivalent to the layers of L1, L2, L3, and L7 of MS-CNN), the channel of convolution kernels in L2 is 10, the size is (57,1); the channel of convolution kernels in L3 is 50, the size is (13,1); the other parameters are the same as MS-CNN.
- CNN-RE (Convolutional Neural Network with Reduced Electrode): This model is identical to CNN-FE except using a small set of electrodes (i.e., FZ, CZ, PZ, P3, P4, PO7, PO8 and Oz) and the kernel size setting of L2 is (8,1).
- CNN-RM (Convolutional Neural Network with Reduced Map): This model is identical to CNN-FE except that it only has one map in the first hidden layer and the channel of kernels in L2 is 1.
- SVM (Support Vector Machine): The data are classified by SVM after preprocessing.
- DCPM (Discriminative Canonical Pattern Matching): DCPM is a robust classification algorithm for evaluating various ERPs. DCPM consists of three major parts: the construction of discriminative spatial patterns (DSPs), the construction of CCA patterns and pattern matching.

B. Subject-unspecific Results

To build the subject-unspecific model, precontest datasets were used as public datasets to train six models (i.e., DCPM, SVM, CNN-FE, CNN-RE CNN-RM and MS-CNN). In Table 2, we listed the accuracies and ITRs of each subject under different numbers of repetitions based on MS-CNN. The accuracy increased with increasing repetition times, while the ITR exhibited an inverse-U shape (first increasing and then decreasing after four repetitions). The averaged accuracy reached its maximum of 58.33% in the case of six repetitions. The maximum of the averaged ITR was 13.49 bits/min in the case of four repetitions, and the ITR decreased to 10.05 bits/min with six repetitions. The respective accuracies and ITRs for each participant can be found in the supplementary materials.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIM.2021.3067943, IEEE Transactions on Instrumentation and Measurement

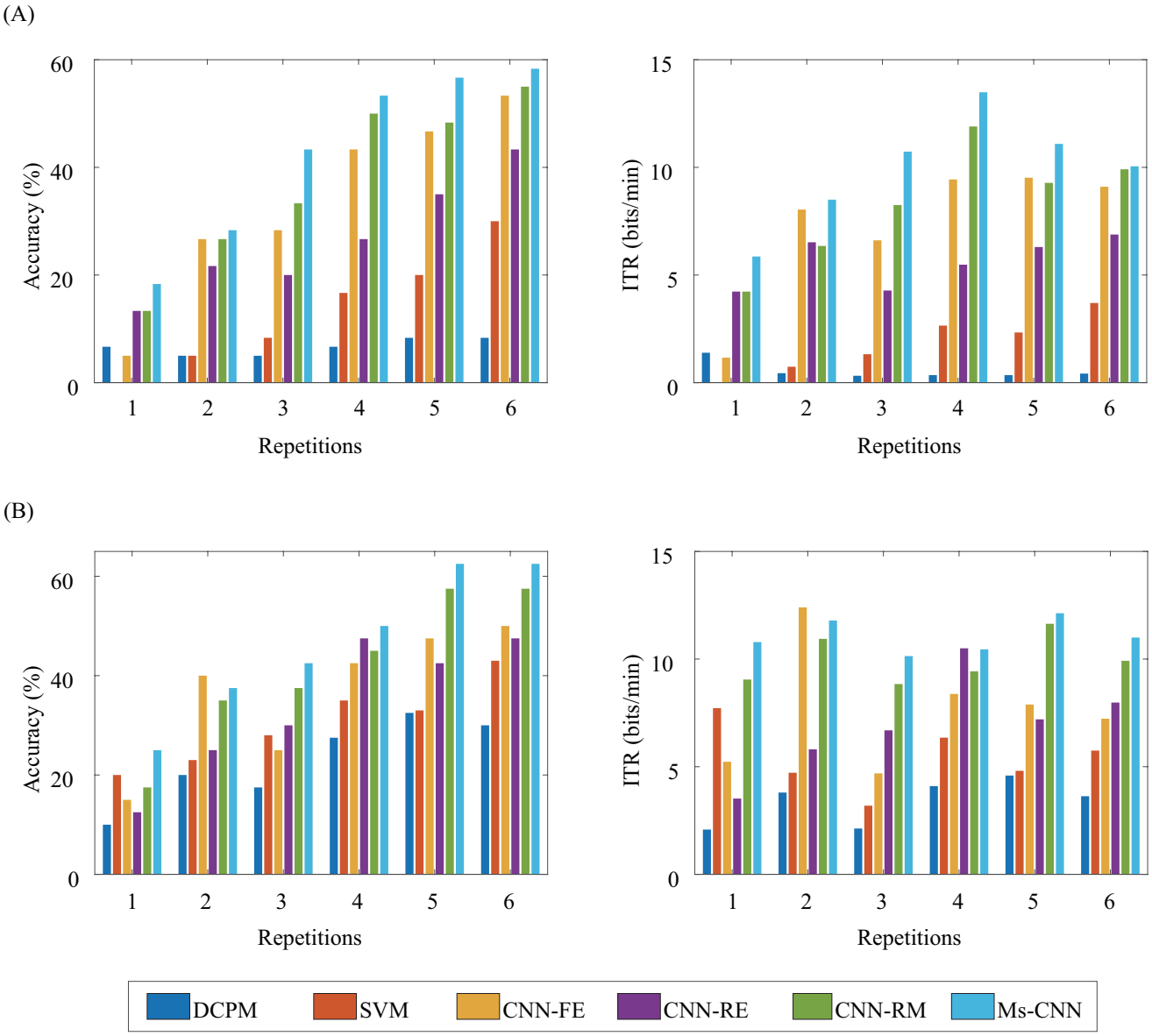

Fig. 2: (A) Averaged accuracies and ITRs obtained by DCPM, SVM, CNN-FE, CNN-RE, CNN-RM and MS-CNN under repetitions from one to six for the subject-unspecific condition. (B) Averaged accuracies and ITRs obtained by DCPM, SVM, CNN-FE, CNN-RE, CNN-RM and MS-CNN under repetitions from one to six for the subject-specific condition.

In our previous studies [50] (ResearchGate available), we compared three shallow classifiers (Fisher LDA, SVM with linear kernel and SVM with Gauss kernel) and found that the shallow classifiers were available with relevant higher accuracies. By analyzing the relationship between the accuracy and the number of repetitions, we found that more than six times of repetitions were required to obtain satisfactory accuracy (above 80%) using shallow classifiers. However, a deep learning model achieved a high accuracy with EEG signals averaged only four times. To evaluate the performance of our proposed MS-CNN model, we compared it with three other widely used deep learning models (i.e., CNN-FE, CNN-RE, and CNN-RM), one shallow classifier (i.e., SVM with Gauss kernel) and one template matching model (i.e., DCPM).

Fig. 2 shows the averaged recognition accuracy of six subjects with different classification methods. The average accuracies under six repetitions were 8.33%, 30%, 53.33%, 43.33%, 55% and 58.33%, while the ITRs were 0.42 bits/min, 3.7 bits/min, 9.1 bits/min, 6.88 bits/min, 9.91 bits/min and 10.05 bits/min, corresponding to DCPM, SVM, CNN-FE, CNN-RE CNN-RM and MS-CNN, respectively. In line with our main finding, the recognition accuracy improved with increasing repetition, while the ITR first increased and then declined. Such a result manifested that we obtained a relatively high ITR with the optimal repetition (four times), although the accuracy was not highest. In addition, because the participants were randomly assigned, the performance of the subjects varied substantially (see the supplementary results for individual performance). In fact, one of the greatest challenges for BCI implementation is 'BCI illiteracy', which refers to a

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIM.2021.3067943, IEEE Transactions on Instrumentation and Measurement

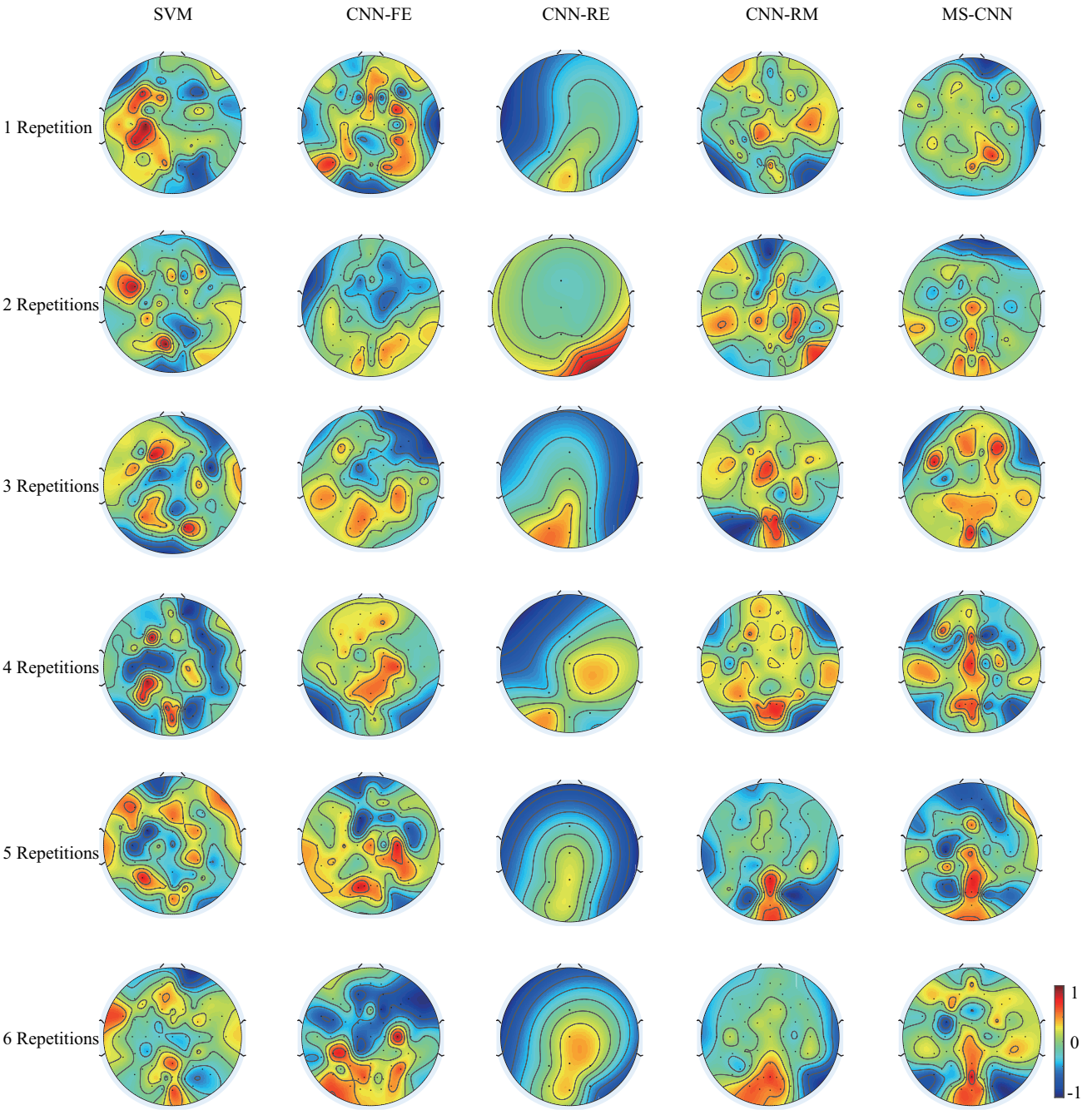

Fig. 3: Weight visualization of each method (repetitions 1-6 from upper to bottom). The weights of the first hidden layer (i.e., Layer L2) were visualized for each deep learning model (i.e., CNN-FE, CNN-RE, CNN-RM and MS-CNN). Each weight corresponds to each electrode (57 in total) used for EEG recording. Similarly, the weights of SVM correspond to electrodes, obtained by averaging across data points (150 in our case).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIM.2021.3067943, IEEE Transactions on Instrumentation and Measurement

phenomenon that occurs in a non-negligible portion of users who cannot properly use a BCI system [51].

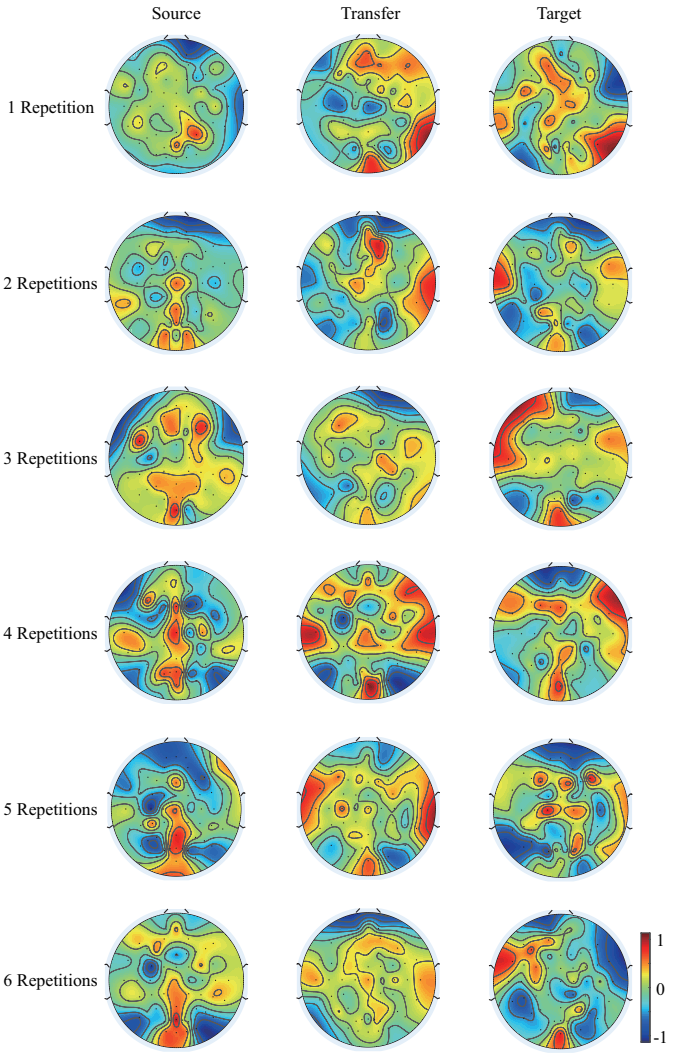

Fig. 4: Weight visualization of MS-CNN (repetitions 1-6 from upper to bottom). The weights of the first hidden layer (i.e., Layer L2) were visualized for MS-CNN from the source domain to target domain (i.e., source: subject-unspecific model; transfer: subject-specific model; and target: subject-specific model without transfer learning). Each weight corresponds to each electrode (57 in total) used for EEG recording. Transfer learning increases the similarity between the source domain data and target domain data.

The visualization of connection weights between layers allows us to inspect the importance of the electrodes. The L2 layer acted as a spatial filter and helped in understanding the importance of the electrodes. The weights from SVM and weights from the first convolution layer of the universal models are displayed in Fig. 3. The red color represents the weights with a high value, which indicates that this electrode has a high discriminant power. With the increase in the number of repetitions, the electrodes became relatively dense in a particular location when the repetition reached four times or more. Furthermore, Cz , Pz , Oz, and POz had higher discriminant powers and showed stability in the MS-CNN, which was in agreement with the results of the literature [35]. Physiologically, it is generally accepted that P300 often occurs primarily in the parietal region and the universal spatial filter in the subject-unspecific group indeed helps to capture P300 potential.

TABLE III: The runtime comparison among deep learning models for the subject-specific condition.

Methods	Runtime (s)	
	Train	Test
MS-CNN	4.441	0.00403
CNN-FE	1.734	0.00359
CNN-RE	2.907	0.00156
CNN-RM	4.914	0.00385

C. Subject-specific Results

EEG signals are highly subject-specific and vary greatly among different individuals [44]. Using subject-specific datasets, we can transfer the individual information to the universal MS-CNN model by transfer learning techniques [49]. During the subject-specific competition, four participants were selected randomly and assigned. The character recognition accuracy and ITR under different numbers of repetitions for four subjects are presented in Table 2. As expected, we found that the recognition accuracy increases with the number of repetitions. However, there are differences between different individuals, and some subjects can obtain high accuracy under few repetitions. Although the ITR can achieve a high value, theoretically, within the first round, it is practically hard to correctly identify P300 using a single trial [52]. Based on the universal models built in the above section, we adapted the model using a transfer learning strategy, where the model parameters were fine-tuned by a small portion of the subject-specific data.

The averaged recognition accuracy and ITR of four subjects using different classifiers are shown in Fig. 2. Similar to the findings in the subject-unspecific group, an increasing trend in recognition accuracy was revealed, with the highest accuracy achieved at six repetitions (i.e., DCPM, 30%; SVM, 43%; CNN-FE, 50%; CNN-RE, 47.5%; CNN-RM, 57.5%; and MS-CNN, 62.5%) and the proposed MS-CNN outperformed the other four methods. In addition, we found that the proposed MS-CNN method also exhibited superior ITR (i.e., DCPM, 3.63 bits/min; SVM, 5.75 bits/min; CNN-FE, 7.23 bits/min; CNN-RE, 7.98 bits/min; CNN-RM, 9.92 bits/min and MS-CNN, 11 bits/min) at the same repetition.

To build the subject-specific model, enhancing the similarity between different individuals is very important. The benefit of transfer learning can be seen in Fig. 4. For P300 detection, both the source domain and target domain have obvious characteristics in the occipital region and parietal lobe region and the weights of the source domain have some characteristic information of the target domain after transfer learning.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIM.2021.3067943, IEEE Transactions on Instrumentation and Measurement

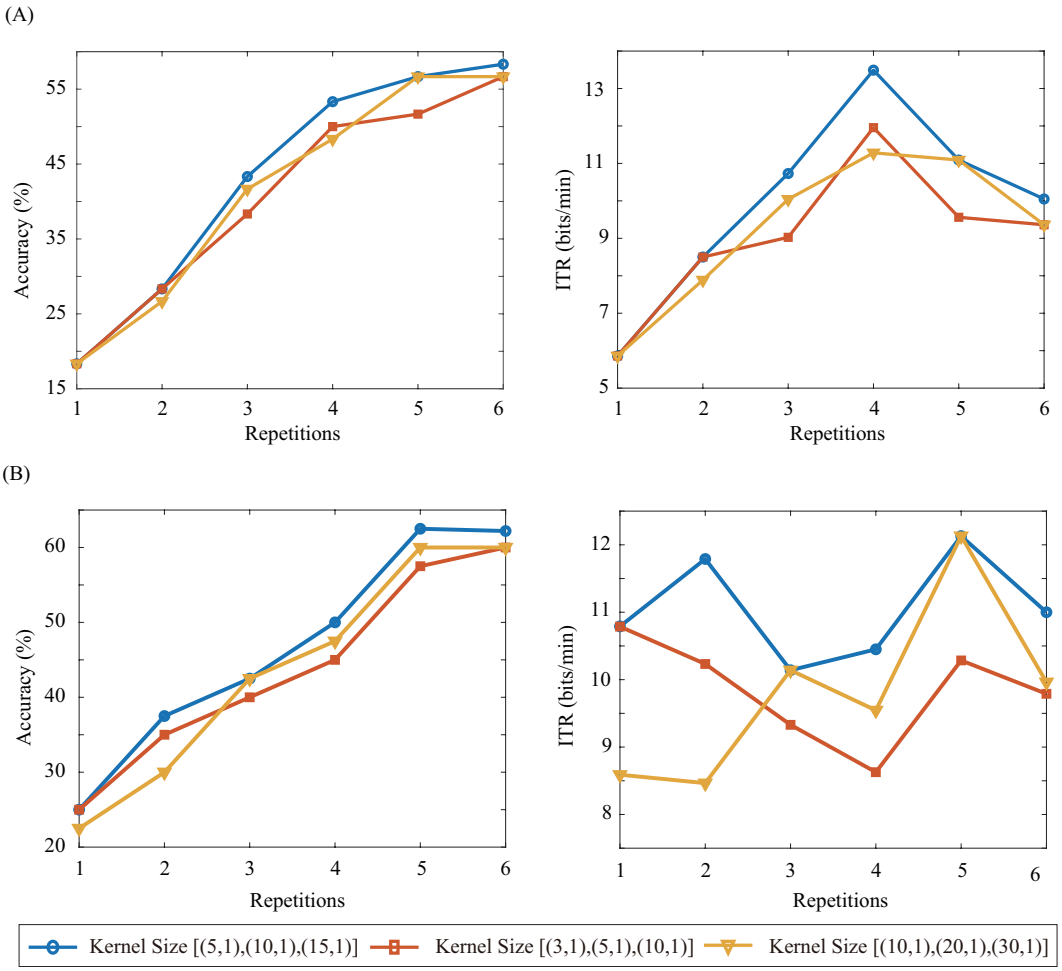

Fig. 5: When the size of convolution kernels of parallel convolution layer is different, the relationship between the correct rate of character recognition and the number of experiments. (A) shows the averaged accuracy and ITR obtained by the MS-CNN with repetitions from one to six in the subject-unspecific group. (B) shows the target detection accuracy and ITR obtained by the MS-CNN methods using one to six repetitions average for the four subjects in subject-specific group.

TABLE IV: Comparison of the averaged ITR with the results obtained in other literatures.

Dataset	Repetitions	Method											
		SVM		ESVM [16], [53]		CNN-FE [35]		CNN-RE [35]		CNN-RM [35]		MS-CNN	
		Acc.	ITR	Acc.	ITR	Acc.	ITR	Acc.	ITR	Acc.	ITR	Acc.	ITR
2003	5	53.5	8.26	73.5	13.74	70	12.69	58.5	10.23	57	9.28	70	**14.26**
2019	5	32.5	4.81	40	5.85	47.5	7.89	42.5	7.20	57.5	11.64	62.5	**12.13**

[1] 2003 denotes the dataset from the 2003 International BCI competition.
[2] 2019 denotes the dataset from the 2019 BCI Controlled Robot Contest.

The runtime comparison among four deep learning models (MS-CNN, CNN-FE, CNN-RE, CNN-RM) in the subject-specific condition is given in Table 3. The training time represents the time required to train the model, and the test time represents the time required to recognize a character. The CNN-FE is the fast technique and the run time for the CNN-RM is the longest. Although the MS-CNN does not have a fast executive speed, the MS-CNN can provide excellent identification performance in comparison with other deep learning models. A compromise between execution efficiency and identification performance indicates that the MS-CNN is still acceptable for real applications.

D. Methodological Considerations

First, we proposed a multi-scale convolutional neural network including three different convolutional kernels [(5,1), (10,1), (15,1)] in the present work. While the other three comparative CNN models include only one kernel, the size is 1×13 for all. To accurately extract the time-varying characteristics of P300, one possible solution is to increase the number of repetitions, and another alliterative solution is to add multi-scale information. Our experimental results showed that multi-scale information can help to improve the generalization performance of the MS-CNN model and highlighted the potential of multi-scale techniques in building

robust features [54].

To select the convolution kernel size, we provided three groups of convolution kernels and their corresponding performances in Fig. 5. The accuracy and ITR of kernel size [(5,1), (10,1), (15,1)] were better than [(3,1), (4,1), (10,1)] and [(10,1), (20,1), (30,1)]. The reason is that the P300 ERP has a certain time range from generation to disappearance. If the convolution kernel is small, the time-domain characteristics obtained by network calculation have little significance, while a larger convolution kernel leads to a rapid increase in computational complexity.

Second, ITR is an important indicator to evaluate the performance of P300-based BCIs. We further compared the performance reported in previous studies [16], [35], [53] based on the open International BCI competition III Data set II [44]. As shown in Table 4, if the number of repetitions is set to five, the highest ITR is 14.26 bits/min, which is achieved by the proposed MS-CNN. Although the subjects from the BCI Controlled Robot Contest of the 2019 World Robot Conference are independent of BCI competition III, the highest ITR of 12.13 bits/min is still achieved by MS-CNN. In summary, our proposed method shows strong generalization performance for different datasets. Compared with SSVEP-based BCI, P300-based BCI is not superior in the higher ITR; however, the greatest advantage is that it does not cause visual fatigue in a long-term usage. In fact, most of P300-based BCIs work in synchronous paradigm mode. In our previous study, we attempted to adjust the number of repetitions dynamically according to the output of SVM to enhance ITR and proposed an asynchronous TV remote control system [50]. In our future work, we will combine a deep learning method and an asynchronous strategy to further enhance the information transmission rate for P300-based BCI.

Finally, most P300-based BCIs are subject-specific and can perform well in P300 detection. However, only a limited number of studies have reported universal models that can realize cross-subject models, making subject-independent P300-based BCI system a challenging problem. The subject-unspecific group of the BCI Controlled Robot Contest at the 2019 World Robot Conference was introduced to promote the potential solution for this challenge. Furthermore, the subjects were randomly recruited and BCI illiteracy was inevitable. All of these factors compromise the practical implementation of BCI techniques [55]. This requires a robust model with high ITR. The proposed MS-CNN showed superiority in terms of the combination of detection accuracy and ITR. In the future, deep learning models can be further improved by considering a few advanced strategies, such as multi-modality, functional connectivity, and tensor deep learning. Detailed considerations can be found in a perspective paper [56].

V. Conclusion

In this study, a multi-scale convolutional neural network (MS-CNN) model was proposed to detect P300. The proposed model was assessed in both subject-unspecific and subject-specific scenarios in the BCI Controlled Robot Contest at the 2019 World Robot Conference. The proposed model achieved the best performance in the contest and we received the champion award in the P300 Competition. The method presented in this paper may pave a promising path for taking a further step towards efficient implementation of P300-based spelling system and help disabled people improve their quality of life.

References

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

[2] J. Li, Y. Liu, Z. Lu, and L. Zhang, "A competitive brain computer interface: Multi-person car racing system," *Conf Proc IEEE Eng Med Biol Soc*, vol. 2013, no. 2013, pp. 2200–2203, 2013.

[3] R. Brice, G. Cuntai, Z. Haihong, W. Chuanchu, T. Cheeleong, H. A. Marcelo, and B. Etienne, "A brain controlled wheelchair to navigate in familiar environments," *Neural Systems and Rehabilitationngineering, IEEE Transactions on*, vol. 18, no. 6, pp. 590–598, 2010.

[4] J. Li, J. Liang, Q. Zhao, J. Li, K. Hong, and L. Zhang, "Design of assistive wheelchair system directly steered by human thoughts," *International journal of neural systems*, vol. 23, no. 03, p. 1350013, 2013.

[5] E. Başar, C. Başar-Eroglu, B. Rosen, and A. Schütt, "A new approach to endogenous event-related potentials in man: relation between eeg and p300-wave," *International Journal of Neuroscience*, vol. 24, no. 1, pp. 1–21, 1984.

[6] A. Pinegger, J. Faller, S. Halder, S. C. Wriessnegger, and G. R. Müller-Putz, "Control or non-control state: that is the question! an asynchronous visual p300-based bci approach," *Journal of Neural Engineering*, vol. 12, no. 1, p. 014001, 2015.

[7] F. Aloise, P. Aricò, F. Schettini, S. Salinari, D. Mattia, and F. Cincotti, "Asynchronous gaze-independent event-related potential-based brain–computer interface," *Artificial Intelligence in Medicine*, vol. 59, no. 2, pp. 61–69, 2013.

[8] B. Wu, Y. Su, J. Zhang, X. Li, J. Zhang, W. Cheng, and X. Zheng, "A virtual chinese keyboard bci system based on p300 potentials," *Acta Electronica Sinica*, vol. 37, no. 8, pp. 1733–1745, 2009.

[9] B. Dal Seno, M. Matteucci, and L. Mainardi, "Online detection of p300 and error potentials in a bci speller," *Computational Intelligence and Neuroscience*, vol. 2010, p. 11, 2010.

[10] N. Haghighatpanah, R. Amirfattahi, V. Abootalebi, and B. Nazari, "A single channel-single trial p300 detection algorithm," *2013 21st Iranian Conference on Electrical Engineering (ICEE)*, pp. 1–5, 2013.

[11] S. Yamaguchi and R. T. Knight, "Age effects on the p300 to novel somatosensory stimuli," *Electroencephalogr Clin Neurophysiol*, vol. 78, no. 4, pp. 297–301, 1991.

[12] J. Polich and A. Kok, "Cognitive and biological determinants of p300: an integrative review," *Biological Psychology*, vol. 41, no. 2, pp. 103–146, 1995.

[13] L. A. Farwell and E. Donchin, "Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials," *Electroencephalography and Clinical Neurophysiology*, vol. 70, no. 6, pp. 510–523, 1988.

[14] Y. Liu, H. Ayaz, A. Curtin, P. A. Shewokis, and B. Onaral, "Detection of attention shift for asynchronous p300-based bci," in *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 2012, pp. 3850–3853.

[15] X. Xiao, M. Xu, J. Jin, Y. Wang, T. P. Jung, and D. Ming, "Discriminative canonical pattern matching for single-trial classification of erp components," *IEEE Transactions on Biomedical Engineering*, vol. 67, no. 8, pp. 2266–2275, 2020.

[16] A. Rakotomamonjy and V. Guigue, "Bci competition iii: dataset ii-ensemble of svms for bci p300 speller," *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 3, pp. 1147–1154, 2008.

[17] Y. Li, Z. Ma, W. Lu, and Y. Li, "Automatic removal of the eye blink artifact from eeg using an ica-based template matching approach," *Physiological Measurement*, vol. 27, no. 4, p. 425, 2006.

[18] D. J. Krusienski, E. W. Sellers, D. J. McFarland, T. M. Vaughan, and J. R. Wolpaw, "Toward enhanced p300 speller performance," *Journal of Neuroscience Methods*, vol. 167, no. 1, pp. 15–21, 2008.

[19] J. Mak, Y. Arbel, J. W. Minett, L. M. McCane, B. Yuksel, D. Ryan, D. Thompson, L. Bianchi, and D. Erdogmus, "Optimizing the p300-based brain–computer interface: current status, limitations and future directions," *Journal of Neural Engineering*, vol. 8, no. 2, p. 025003, 2011.

[20] H. Cecotti, B. Rivet, M. Congedo, C. Jutten, O. Bertrand, E. Maby, and J. Mattout, "A robust sensor-selection method for p300 brain–computer interfaces," *Journal of Neural Engineering*, vol. 8, no. 1, p. 016001, 2011.

[21] M. Xu, H. Qi, L. Ma, C. Sun, L. Zhang, B. Wan, T. Yin, and D. Ming, "Channel selection based on phase measurement in p300-based brain-computer interface," *PloS One*, vol. 8, no. 4, p. e60608, 2013.

[22] J. Lu, K. Xie, and D. J. McFarland, "Adaptive spatio-temporal filtering for movement related potentials in eeg-based brain–computer interfaces," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, no. 4, pp. 847–857, 2014.

[23] M. Kaper, P. Meinicke, U. Grossekathoefer, T. Lingner, and H. Ritter, "Bci competition 2003-data set iib: support vector machines for the p300 speller paradigm," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 6, pp. 1073–1076, 2004.

[24] J. Li, H. Huai, J. Gao, D. Kong, and L. Wang, "Spatial-temporal dynamic hand gesture recognition via hybrid deep learning model," *Journal on Multimodal User Interfaces*, pp. 1–9, 2019.

[25] P. Wang, A. Jiang, X. Liu, J. Shang, and L. Zhang, "Lstm-based eeg classification in motor imagery tasks," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 26, no. 11, pp. 2086–2095, 2018.

[26] M. Li, W. Zhu, M. Zhang, Y. Sun, and Z. Wang, "The novel recognition method with optimal wavelet packet and lstm based recurrent neural network," *2017 IEEE International Conference on Mechatronics and Automation (ICMA)*, pp. 584–589, 2017.

[27] Y. LeCun, L. Bottou, Y. Bengio, P. Haffner *et al.*, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.

[28] S. Min, L. Byunghan, and Y. Sungroh, "Deep learning in bioinformatics," *Briefings in Bioinformatics*, no. 5, p. 851, 2017.

[29] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241.

[30] C. Farabet, C. Couprie, L. Najman, and Y. LeCun, "Learning hierarchical features for scene labeling," *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 8, pp. 1915–1929, 2012.

[31] A. Sharif Razavian, H. Azizpour, J. Sullivan, and S. Carlsson, "Cnn features off-the-shelf: an astounding baseline for recognition," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 806–813, 2014.

[32] A. J. X. Guo and F. Zhu, "A cnn-based spatial feature fusion algorithm for hyperspectral imagery classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 9, pp. 7170–7181, 2019.

[33] S. Hwang, K. Hong, G. Son, and H. Byun, "Learning cnn features from de features for eeg-based emotion recognition," *Pattern Analysis and Applications*, pp. 1–13, 2019.

[34] X. Tang, J. Yang, and H. Wan, "A hybrid sae and cnn classifier for motor imagery eeg classification," *Computer Science On-line Conference*, pp. 265–278, 2018.

[35] H. Cecotti and A. Graser, "Convolutional neural networks for p300 detection with application to brain-computer interfaces," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 3, pp. 433–445, 2010.

[36] H. Wang, L. Xu, A. Bezerianos, C. Chen, and Z. Zhang, "Linking attention-based multiscale cnn with dynamical gcn for driving fatigue detection," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–11, 2021.

[37] T. Wang, A. Bezerianos, A. Cichocki, and J. Li, "Multi-kernel capsule network for schizophrenia identification," *IEEE Transactions on Cybernetics*, p. DOI: 10.1109/TCYB.2020.3035282, 2019.

[38] G. Zhu, Y. Li, P. P. Wen, S. Wang, and N. Zhong, "Unsupervised classification of epileptic eeg signals with multi scale k-means algorithm," pp. 158–167, 2013.

[39] S. Gao, M. M. Cheng, K. Zhao, X. Y. Zhang, and P. H. S. Torr, "Res2net: A new multi-scale backbone architecture," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PP, no. 99, pp. 1–1, 2019.

[40] Y. Chen, H. Fan, B. Xu, Z. Yan, Y. Kalantidis, M. Rohrbach, Y. Shuicheng, and J. Feng, "Drop an octave: Reducing spatial redundancy in convolutional neural networks with octave convolution," *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, 2020.

[41] Y. Li, Z. Kuang, Y. Chen, and W. Zhang, "Data-driven neuron allocation for scale aggregation networks," *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 11 518–11 526, 2019.

[42] M. Tan and Q. V. Le, "Mixconv: Mixed depthwise convolutional kernels," *arXiv:1907.09595*, 2019.

[43] I. C. Duta, L. Liu, F. Zhu, and L. Shao, "Pyramidal convolution: Rethinking convolutional neural networks for visual recognition," *arXiv:2006.11538*, 2020.

[44] B. Blankertz, K.-R. Muller, D. J. Krusienski, G. Schalk, J. R. Wolpaw, A. Schlogl, G. Pfurtscheller, J. R. Millan, M. Schroder, and N. Birbaumer, "The bci competition iii: Validating alternative approaches to actual bci problems," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 14, no. 2, pp. 153–159, 2006.

[45] B. Blankertz, M. Krauledat, G. Dornhege, J. Williamson, R. Murray-Smith, and K.-R. Müller, "A note on brain actuated spelling with the berlin brain-computer interface," *International Conference on Universal Access in Human-Computer Interaction*, pp. 759–768, 2007.

[46] K. Yu, T. R. Lin, H. Ma, X. Li, and X. Li, "A multi-stage semi-supervised learning approach for intelligent fault diagnosis of rolling bearing using data augmentation and metric learning," *Mechanical Systems and Signal Processing*, vol. 146, p. 107043, 2021.

[47] K. Yu, H. Ma, T. Lin, and X. Li, "A consistency regularization based semi-supervised learning approach for intelligent fault diagnosis of rolling bearing," *Measurement*, vol. 165, p. 107987, 2020.

[48] K. Yu, Q. Fu, H. Ma, and T. Lin, "Simulation data driven weakly supervised adversarial domain adaptation approach for intelligent cross-machine fault diagnosis," *Structural Health Monitoring*, 2021.

[49] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on knowledge and data engineering*, vol. 22, no. 10, pp. 1345–1359, 2009.

[50] H.-T. Wang and H.-L. Zou, "Asynchronous tv remote control system based on event-related potential brain-computer interface," *Control Theory & Applications*, vol. 29, no. 11, pp. 1507–1511, 2012.

[51] C. Vidaurre and B. Blankertz, "Towards a cure for bci illiteracy," *Brain Topography*, vol. 23, no. 2, pp. 194–198, 2009.

[52] Z. Lin, Y. Zeng, L. Tong, H. Zhang, C. Zhang, and B. Yan, "Method for enhancing single-trial p300 detection by introducing the complexity degree of image information in rapid serial visual presentation tasks," *PloS One*, vol. 12, no. 12, p. e0184713, 2017.

[53] "Bci competition iii final results, http://ida.first.fraunhofer.de/projects/bci/competition-iii/results/, 2008."

[54] J. Wang, J. Zhuang, L. Duan, and W. Cheng, "A multi-scale convolution neural network for featureless fault diagnosis," in *2016 International Symposium on Flexible Automation (ISFA)*. IEEE, 2016, pp. 65–70.

[55] R. Carabalona, "The role of the interplay between stimulus type and timing in explaining bci-illiteracy for visual p300-based brain-computer interfaces," *Frontiers in Neuroscience*, vol. 11, p. 363, 2017.

[56] J. Li, "Thoughts on neurophysiological signal analysis and classification," *Brain Science Advances*, vol. 6, no. 3, pp. 210–223, 2020.