

Research Article

A Collaborative Brain-Computer Interface Framework for Enhancing Group Detection Performance of Dynamic Visual Targets

Xiyu Song ¹, Ying Zeng ^{1,2}, Li Tong ¹, Jun Shu ¹, Qiang Yang ³, Jian Kou ⁴,
Minghua Sun ⁵ and Bin Yan ¹

¹Henan Key Laboratory of Imaging and Intelligent Processing, PLA Strategic Support Force Information Engineering University, Zhengzhou, China

²The Clinical Hospital of Chengdu Brain Science Institute, MOE Key Lab for Neuro Information, University of Electronic Science and Technology of China, Chengdu, China

³Xi'an Satellite Control Center, Hangzhou, China

⁴PLA 32317 Force, Wulumuqi, China

⁵Department of Radiology, Henan Provincial People's Hospital, Department of Radiology, Central China Fuwai Hospital, Zhengzhou University, Zhengzhou, China

Correspondence should be addressed to Bin Yan; ybspace@hotmail.com

Received 20 October 2021; Revised 12 December 2021; Accepted 14 December 2021; Published 18 January 2022

Academic Editor: Ivan Volosyak

Copyright © 2022 Xiyu Song et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The superiority of collaborative brain-computer interface (cBCI) in performance enhancement makes it an effective way to break through the performance bottleneck of the BCI-based dynamic visual target detection. However, the existing cBCIs focus on multi-mind information fusion with a static and unidirectional mode, lacking the information interaction and learning guidance among multiple agents. Here, we propose a novel cBCI framework to enhance the group detection performance of dynamic visual targets. Specifically, a mutual learning domain adaptation network (MLDANet) with information interaction, dynamic learning, and individual transferring abilities is developed as the core of the cBCI framework. MLDANet takes P3-sSDA network as individual network unit, introduces mutual learning strategy, and establishes a dynamic interactive learning mechanism between individual networks and collaborative decision-making at the neural decision level. The results indicate that the proposed MLDANet-cBCI framework can achieve the best group detection performance, and the mutual learning strategy can improve the detection ability of individual networks. In MLDANet-cBCI, the F1 scores of collaborative detection and individual network are 0.12 and 0.19 higher than those in the multi-classifier cBCI, respectively, when three minds collaborate. Thus, the proposed framework breaks through the traditional multi-mind collaborative mode and exhibits a superior group detection performance of dynamic visual targets, which is also of great significance for the practical application of multi-mind collaboration.

1. Introduction

Brain-computer interface (BCI) technology aims to build an interaction bridge between human and computer and provide a new technical means for the brain to control and monitor the external environment. Advanced BCI technology can not only help to improve the movement abilities of patients with physical disorders [1, 2], but also enhance such abilities of healthy people [3–7]. Affected by changes in

the surrounding environment and in the psychological factors of users, a single-mind BCI shows limited performance and thus is hard to translate into practical application. Multi-mind collaborative brain-computer interfaces (cBCIs) have special advantages in enhancing the group detection performance. Multi-mind BCIs are equivalent to multiple information processing systems, which show higher group decision-making performance and stronger robustness. In addition, multi-mind collaborative work is

more conducive to the future development of human-computer interaction socialization. P300-based BCIs broaden the BCI's practical application. The classical visual speller and target detection are based on P300 identification [8, 9]. For the task of dynamic target detection, the dynamics of video background, the uncertainty of distractors, and the jitter of detection latency increase the detection difficulty, resulting in the limitations of single-mind BCI [10, 11]. The cBCI can be considered as a good strategy to solve the problem, which will contribute to improving the stability and accuracy of comprehensive discrimination [12, 13]. Therefore, building a novel cBCI framework to highlight the performance advantages of multi-mind enhancement has become the research focus to improve the performance of dynamic visual target detection.

The cBCIs are praised as one of the most promising applications in human augmentation [14–16]. The group decision-making capability can be improved by integrating the multi-mind information and optimizing collaborative strategies [17, 18]. For the task of target detection, the multi-mind collaborative information integration mainly includes three levels: signal-level fusion, feature-level fusion, and decision-level fusion. Signal-level fusion is the simplest way to improve the signal-to-noise ratio (SNR) of EEG signals, where multi-participant EEG signals are averaged and input to a classifier. Feature-level fusion is the classification of the averaged or concatenated features from multi-participant EEG signals. Both signal-level fusion and feature-level fusion belong to single-classifier cBCI (SC-cBCI). Decision fusion merges the multiple classifiers' decision-making results into the final decision. Each classifier corresponds to one participant. Thus, decision fusion is also known as multi-classifier cBCI (MC-cBCI). The specific decision emergence strategies include averaged decision, weighted decision, and majority voting on the decision probability layer. The cBCI has attracted scholars' interest in target detection. To explore the best fusion level, Matran-Fernandez et al. preliminarily verified that the decision fusion in the cBCI performs better than any single-mind BCI (sBCI) for the single-trial P300 detection [19–21]. The relevant studies indicated that the decision-level fusion performs better than signal-level fusion and feature-level fusion in the cBCI [22]. To explore the best decision emergence strategies, Cecotti et al. [23, 24] found that the averaged decision performs better than weighted decision and voting strategy; Davide et al. [15, 25–27], Yuan et al. [28, 29], and Jiang et al. [30] trained individual decision weights through least angle regression (LARS) method, two-layer SVM, and a combination of SVM and LDA classifiers to improve group detection performance. Since the above findings are inconsistent, the selection of decision fusion strategies relies on the specific experimental tasks. To introduce the information interaction among multiple minds, Davide et al. [16] studied the impact of individual behavior decision sharing on collaborative behavior decision-making and found that information interaction in the experimental process will lead to the decline of behavior-level collaborative decision-making performance. To fuse more information and improve the group detection performance, Zhang et al. [31] proposed a dual brain collaborative target

detection model, which integrates data fusion and feature fusion to ensure that important information is not missed; Eckstein et al. [32] explored and compared the impact of the user number on the collaborative decision performance and found that the best collaborative detection performance generally requires 5~10 users. These studies provide technical reference and theoretical support for the design of multi-mind collaborative experimental paradigm and the development of cBCI framework.

Studies on multi-mind collaborative target detection have achieved remarkable results. However, some concerning issues remain. Firstly, in the current cBCIs, the computing models are static and unidirectional [18, 32–34] only be known as “multi-mind fusion” rather than “multi-mind collaboration.” Real collaboration should involve multi-mind information interaction, which is a dynamic learning process. Given the higher error rate caused by the individual communication in the experimental process [16], relative to behavior level, the information interaction can be established at the neural decision layer. Thus, a dynamic interactive cBCI framework at neural decision layer could be considered to improve the performance. Secondly, previous studies required participants to increase the preparation time to collect labeled signals for an individual-specific computing model [26, 27, 34]. For the dynamic visual target detection, unsupervised domain adaptation networks, P3-MSDA and P3-sSDA, have been developed as an individual-general network with reliable performance in previous studies [35]. Therefore, it is necessary to develop a novel cBCI framework with information interaction, dynamic learning, and individual transferring abilities to enhance the group detection performance video targets.

In this study, we designed a novel multi-mind cBCI framework based on a mutual learning domain adaptation network (MLDANet), aiming at enhancing the group detection performance of dynamic visual targets. In the framework, a multi-mind synchronous cBCI experimental paradigm for UAV-video vehicle detection is designed; MLDANet is established where the P3-sSDA network is used as the individual network unit. This work made the following contributions.

- (1) The MLDANet-cBCI framework was proposed for achieving better group detection performance. In particular, MLDANet establishes the information interaction and dynamic learning mechanism between individual networks and collaborative decision-making by introducing the mutual learning strategy at the neural decision layer.
- (2) In the MLDANet-cBCI framework, the mutual learning strategy can effectively improve the individual network capability.

2. Materials and Methods

This collaborative brain-computer interface (cBCI) framework is designed to detect dynamic visual targets, as shown in Figure 1. The framework consists of four modules, which are stimulus presentation, synchronous acquisition, data

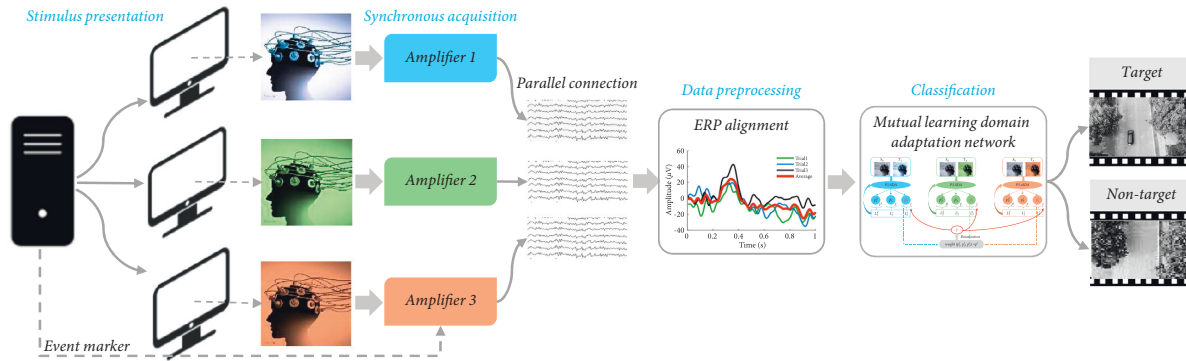


FIGURE 1: Collaborative BCI framework for detecting dynamic visual targets.

preprocessing, and classification. The stimulus presentation module synchronously shows unmanned aerial vehicle (UAV) videos to all participants to detect the vehicles from these videos. The synchronous acquisition module collects multi-mind EEG signals with time synchronization. The data preprocessing module aims to obtain artifact-free and filtered EEG epochs for target and nontarget trials. The classification module is the core of the cBCI framework, where a mutual learning domain adaptation network (MLDANet) is proposed to improve the group detection performance.

2.1. Stimulus Presentation. The experimental paradigm for vehicle detection from UAV videos, reported in our previous study [5], is depicted in Figure 2. The UAV videos recorded traffic conditions while flying along campus streets. A series of video clips were segmented from original videos to construct a stimulus library. One hundred video clips with vehicles (one vehicle per video) and 100 video clips without vehicles were, respectively, regarded as target videos and nontarget videos. In this experiment, the total duration time of videos is about 28 minutes. To alleviate the vision load, we divided all the video clips to 10 blocks and set break time between blocks. In each block, 10 target videos and 10 nontarget videos were randomly presented to the participants in each block. The length of the video clip varied from 4 s to 10 s, and there were 2 s “+” before each video to help participants focus their attention. For each target video, the vehicle could enter the visual field from any direction at any time 1 s after the video stimulus was presented. To overcome the influence of video color and eye movement on visual perception, the video clips ($1920 \times 1080 \text{ px}^2$) were transformed into black and white and reduced to 40% ($768 \times 432 \text{ px}^2$) on the screen center against black background. In particular, the break duration totally depended on the subjects for sufficient relax. On average, the experimental duration time of 10 blocks (including break duration) was around 50 minutes for one participant.

2.2. Synchronous Acquisition. A total of 89 healthy college participants volunteered for this study, with a median age of 25 years (right-handed), all of whom reported normal or corrected-to-normal vision and presented no neurological

problems. They all signed the informed consent form before the experiment. All tests involving human participants were approved by the Ethics Committee of Henan Province People’s Hospital.

In this study, the EEG signals were collected by the g.USBamp (g.tec, Austria) EEG recording with 16 electrodes. The electrode distribution followed the international 10–20 electrode location. The EEG online sampling rate was 600 Hz with band-pass filtering at 0.01–100 Hz and notch filtering at 50 Hz.

The study was comprised of two parts: single-mind experiment and multi-mind experiment. 29 participants were recruited for the single-mind experiment, and each time only a single participant was invited to perform the detection task. The EEG signals collected from the single-mind experiment were used as the training set. In the multi-mind experiment, 20 groups (3 participants in each group) were recruited to perform the detection task together. The acquisition environment for the multi-mind synchronous experiment is shown in Figure 3. The same stimulus materials were simultaneously displayed on three copied displays. There was no communication between the three participants during the experiment. Meanwhile, EEG signals from 3 participants were synchronously collected by three parallel EEG amplifiers and recorded by recording software (g.Recorder). The recording software arranged these 16-channel signals. Channels 1~16, channels 17~32, and channels 33~48 were, respectively, from participant 1, participant 2, and participant 3. In this manner, the synchronization of time, space, and surrounding environment ensured that the influence of external factors on each participant was identical. All the collected data will be made available to peers for any relevant future work.

2.3. Data Preprocessing. The parameter settings of data preprocessing were consistent between single-mind and multi-mind experiments. Firstly, using the fast ICA algorithm and EEGLAB toolbox, the electrooculogram was removed from the original signals. Next, data were filtered to 0.1–10 Hz and downsampled to 100 Hz. Then, target segments and nontarget segments were, respectively, extracted from target video-induced and nontarget video-induced EEG signals. The target trials were segmented starting from

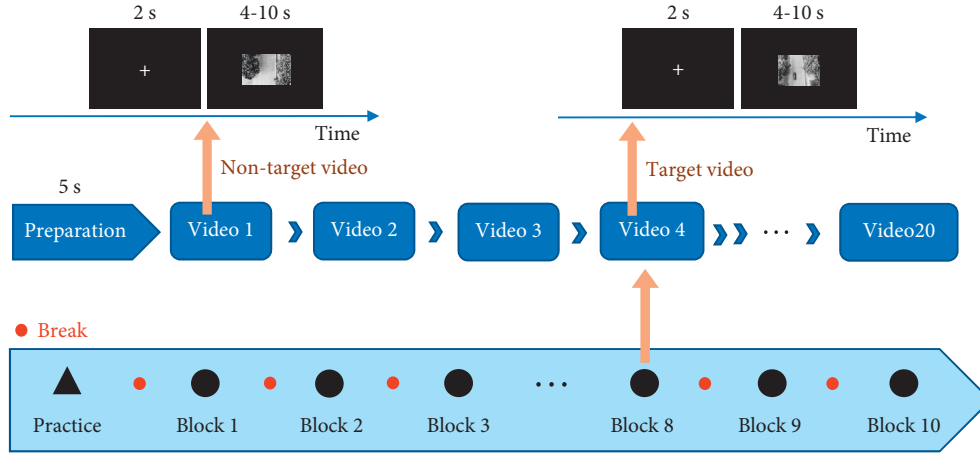


FIGURE 2: Experimental paradigm for vehicle detection in UAV video.



FIGURE 3: The EEG acquisition environment for the multi-mind synchronous experiment.

target onset time. One target video can only induce one target trial, and nontarget trials were segmented from nontarget video-induced signals without overlapping, where one nontarget video can contain several nontarget trials. Thus, 100 target trials from 100 target videos and 521 nontarget trials from 100 nontarget videos were extracted. The signals for each trial were 1500 ms. The size of the single-trial sample was 16×150 (channels \times time sample points).

Since a domain adaptation network was applied to predict the detection performance. Single-mind signals and multi-mind signals were, respectively, used to construct the source domain and the target domain.

2.3.1. Single-Mind Signals (Source Domain). Single-mind signals from 29 participants were employed to construct the training set. In our previous studies [35], individuals with strong P3 responses by P3 map-clustering method as source domains help to achieve better detection performance. Here, the P3 map-clustering method was applied to select excellent individuals as the source domain for better detection

performance. Due to the serious time jitter of P300 latency for the video-induced EEG signals, before using the P3 map-clustering method, individual P3 maps had to be extracted by the event-related potential (ERP) alignment method [11]. The principle of this method is to reduce the spatial dimension of the single-trial signal to construct a one-dimensional target ERP template by the common spatial patterns (CSP) [6, 36], and match all the one-dimensional time series with the ERP template to obtain the aligned P300 signals. Using the constructed 1000 ms ERP template, the size of the aligned trials was determined as 16×100 (channels \times time sample points). The brain topographic map at a peak time of the P300 component was extracted as an individual P3 map. Here, 29 individual P3 maps were obtained from 29 participants. Using a *K-means* distance-clustering method with two clustering centers, 29 maps were clustered into two groups with strong P3 maps and weak P3 maps [35]. Subsequently, the individuals in the strong P3 map group were considered to act as the source domain.

2.3.2. Multi-Mind Signals (Target Domain). Multi-mind signals were used as the testing set. Using the ERP template constructed from the source domain, there were 621 aligned trials in total (100 target trials and 521 nontarget trials) available for each individual, which constituted an imbalanced dataset. The size of the single-trial sample was 16×100 (channels \times time sample points). The validity of the single-trial signals was tested to screen some samples as target domain for domain adaptation network. The threshold value method was adopted for the sample screening. The single-trial signals with maximum amplitude values within $\pm 120 \mu\text{V}$ were regarded as valid signals; otherwise, they were regarded as invalid signals. Thus, each single-trial EEG signal corresponded to two labels (category label and validity label) for each participant.

2.4. Classification Model. Aimed at enhancing the performance of group detection of dynamic visual targets, MLDANet is proposed as shown in Figure 4. The core of MLDANet is to establish the mechanisms of multi-mind

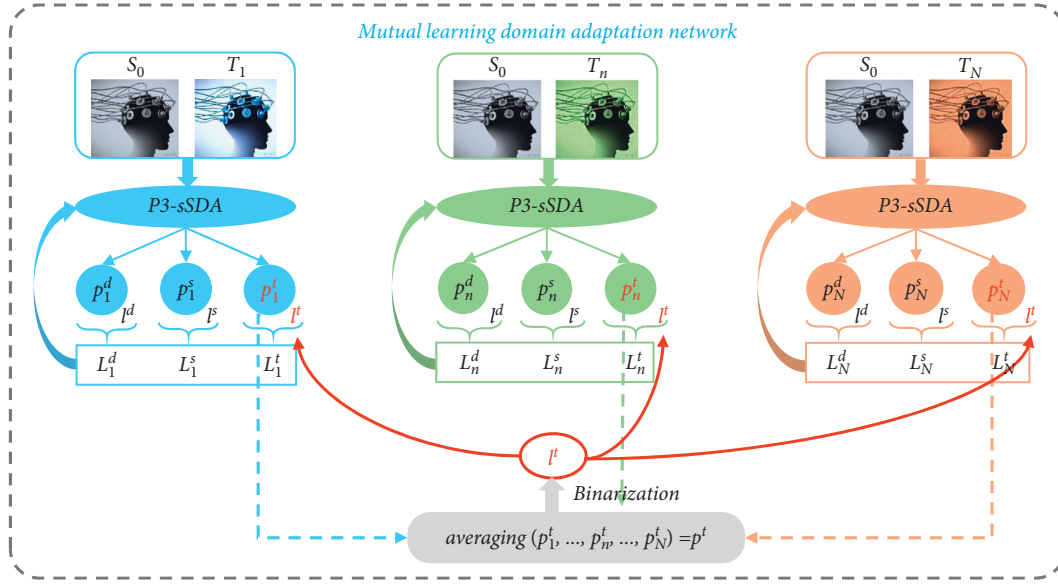


FIGURE 4: The architecture of MLDANet for group detection.

information interaction and dynamic learning at the neural decision layer. In the MLDANet, the collaborative decision-making guides the individual network to re-decision-making for the enhancement of collaborative decision-making performance.

In the proposed framework, an unsupervised single-source domain adaptation network with strong P3 map individuals as the source domain for dynamic visual target detection (P3-sSDA network) is used as the individual network unit [35]. The P3-sSDA network is an individual-generalized model with good performance in EEG-based dynamic visual target detection, as shown in Figure 5. P3-sSDA consists of five parts: source domain selector, feature extractor, domain discriminator, category classifier, and target domain sample selector. In P3-sSDA, a P3 map-clustering method selected the individuals with strong P3 maps as one source domain. Feature extractor extracts the EEG features from video targets-induced EEG signals. Domain discriminator performs the adversarial domain adaptation to eliminate individual differences. Category classifier classifies the EEG features to distinguish target samples and nontarget samples. The testing samples are ranked according to the probability value predicted as target samples. Target domain sample selector selects the samples most like the target samples from testing samples as the target domain samples for the imbalanced data classification. The proportion of samples selected is 80%. In this study, the training individuals and testing individuals were completely independent. Thus, the P3-sSDA network was suitable for establishing individual-generalized cBCI frameworks for dynamic visual target detection. The detailed network architecture was given in our previous work [35].

In the MLDANet, there were N target domain individuals from one group with N P3-sSDA networks, which were used as classifiers that synchronously worked on predicting the binary classification probability of single-

trial signals. Collaborative decision-making was achieved by the decision-making fusion. The mutual learning strategy was introduced between N P3-sSDA networks and collaborative decision-making. Data from the common source domain and different target domains were, respectively, denoted as $S_0, T_1, T_2, \dots, T_N$, where N P3-sSDA networks shared the same source samples and used N different target domain samples. For a single P3-sSDA network, the input sample number of source domain or target domain was named as batch size m . For the n -th target domain individual, m source domain samples $x_1^s, x_2^s, \dots, x_m^s$ and m target domain samples $x_1^{t_n}, x_2^{t_n}, \dots, x_m^{t_n}$ were input into the n -th P3-sSDA network in each batch. Importantly, target domain samples from N target domain individuals were synchronously recorded for the same stimulus scene. Thus, the prediction category labels of the target domain could be shared among N P3-sSDA networks, which was crucial to achieve the information interaction. For each iteration, the P3-sSDA network could output the domain discriminant probability between source domain samples and target domain samples, the category prediction probability of source domain samples, and the category prediction probability of target domain samples. These probabilities were, respectively, denoted as p_n^d, p_n^s, p_n^t for the n -th P3-sSDA network. Since N individuals synchronously received the same stimuli information, the prediction probability, $p_1^t, p_2^t, \dots, p_N^t$, could reflect the discrimination level for the same information from different individuals. In the process of domain adaptation, the category labels of the source domain l^s and the domain labels l^d were available; hence, the category loss of the source domain was given as

$$L_{\text{class}}^{s_{t_n}} = - \sum_{k=1}^m (l_k^s \log p_n^{s,k}), \quad (1)$$

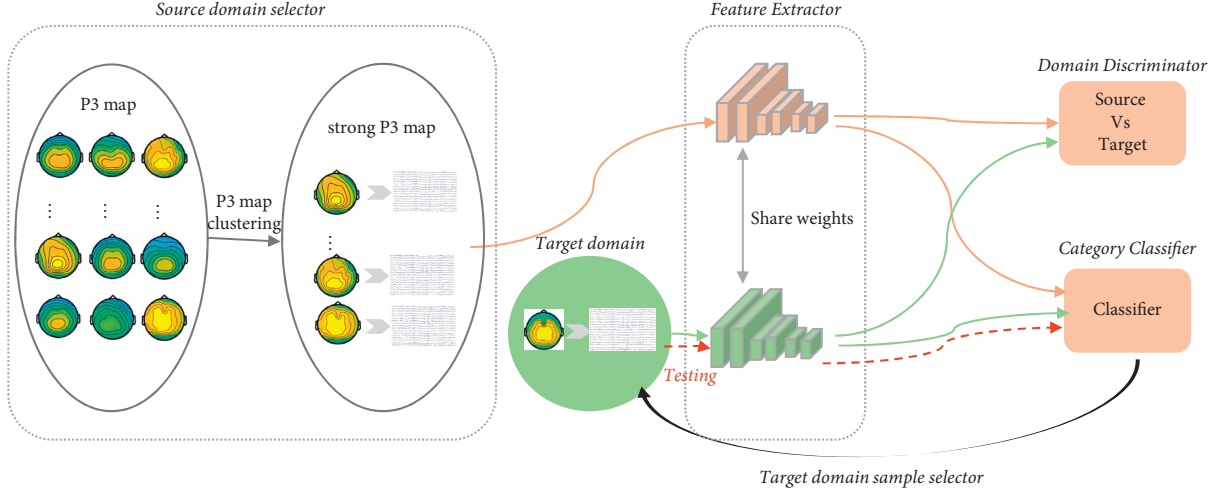


FIGURE 5: P3-sSDA network architecture.

and the domain loss between the source domain and the n -th target domain, $L_{a\ dv}^{s,t_n}$, could be obtained for the n -th target domain:

$$L_{a\ dv}^{s,t_n} = - \sum_{k=1}^m (l_k^{d,s} \log p_n^{d,s,k} + l_k^{d,t} \log p_n^{d,t,k}). \quad (2)$$

Nevertheless, the category loss of the target domain was unknown due to the lack of category labels of the target domain l^t , which would be estimated in each iteration. By averaging the individual predictions, the integrated prediction probability p^t was binarized as l^t :

$$l^t = \text{binary}(p^t) = \text{binary}(\text{averaging}(p_1^t, p_2^t, \dots, p_N^t)), \quad (3)$$

which was viewed as the prediction results of the multiple minds. Moreover, the integrated prediction label l^t was used as the pseudo-label and synchronously fed back to N P3-sSDA networks to calculate the category losses of target domain $\{L_{\text{class}}^t\}_{n=1}^N$.

$$L_{\text{class}}^t = - \sum_{k=1}^m (l_k^t \log p_n^{t,k}). \quad (4)$$

Thus, the category and discrimination losses could be calculated. For the MLDANet with N P3-sSDA network, the entire adversarial learning problem could be described as follows:

$$\min \sum_{n=1}^N (a \times L_{a\ dv}^{s,t_n} + \gamma \times L_{\text{class}}^{s_n} + \beta \times L_{\text{class}}^t), \quad (5)$$

where $L_{a\ dv}^{s,t_n}$, $L_{\text{class}}^{s_n}$, and L_{class}^t denote the domain discrimination loss between the source domain and the n -th target domain, the category loss of the source domain when adapted to the n -th target domain, and the category loss of the n -th target domain, respectively; α , γ , and β are hyperparameters which, respectively, denote discrimination loss weight, category loss weight of source domain, and category loss weight of target domain.

The mechanism of information interaction among multi-mind signals was established through information integration and feedback. In the MLDANet framework, each individual network not only receives the supervision from individual network labels, l^d and l^s , but also refers to the collaborative label l^t , which is calculated from all individual networks. By the backpropagation of collaborative label l^t , all the individual networks can learn from each other and make common progress. This process can help small network training to be more powerful. Different from classical cBCI with once collaborative decision-making, with the iteration and updating of network parameters, the dynamic learning ability of the individual network was established in the MLDANet. Thus, the single P3-sSDA network (single individual) could learn from the source domain, target domain, and collaborative decision-making (group). After the training process of DA, the online testing can be conducted as the procedure of red dotted line in Figure 5. The new testing samples can directly be tested by feature extractor and category classifier.

3. Results

3.1. Source Domain Individuals. A total of 29 individual P3 maps were clustered into a strong P3 group $\{sub6, sub11, sub12, sub13, sub14, sub16, sub17, sub19, sub20, sub21, sub23, sub26, sub28\}$ and a weak P3 group $\{sub1, sub2, sub3, sub4, sub5, sub7, sub8, sub9, sub10, sub15, sub18, sub22, sub24, sub25, sub27, sub29\}$. The topographies of the strong P3 group and the weak P3 group at the P3 peak value are presented in Figure 6, where the parieto-occipital region is closely related to P3 responses. Since studies indicated that individuals with strong P3 maps are more suitable as the source domain, we selected all individuals from the strong P3 group as the source domain. The averaged ERP responses of source domain individuals (13 individuals from the strong P3 group) are shown in Figure 7.

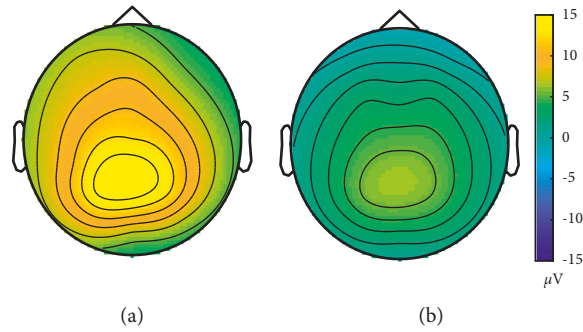


FIGURE 6: The averaged P3 map of two groups: (a) strong P3 map; (b) weak P3 map.

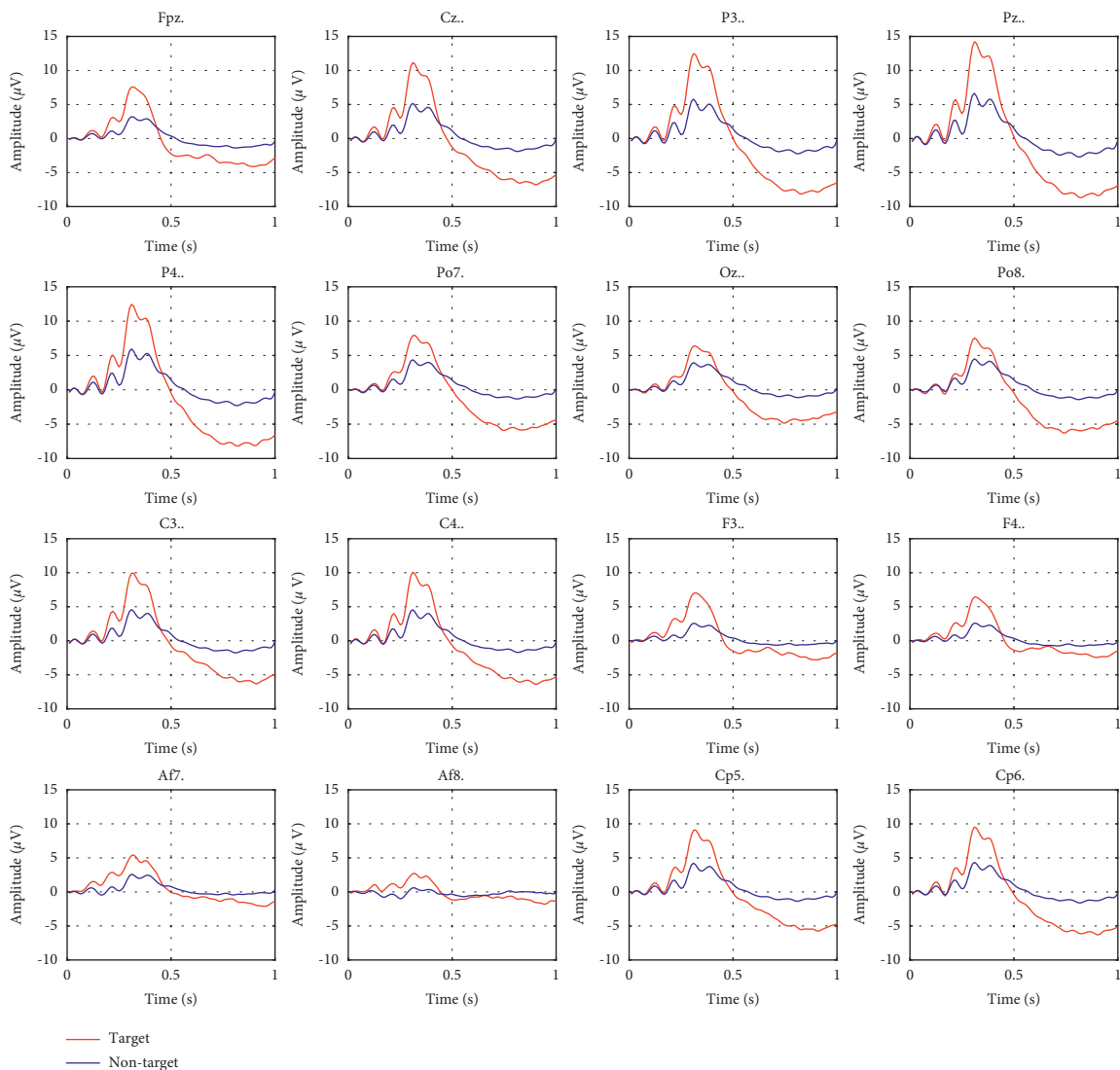


FIGURE 7: The averaged ERP responses of source domain individuals.

3.2. *Detection Performance.* In this work, we compared the detection performances of four BCI frameworks, namely, the single-mind BCI (sBCI), the single-classifier cBCI (SC-cBCI), the multi-classifier cBCI (MC-cBCI), and the proposed MLDANet-cBCI framework. The summary of these BCI frameworks is shown in Figure 8. One individual

participates in the sBCI framework, where the results are averaged from 60 individuals. The SC-cBCI, MC-cBCI, and MLDANet-cBCI frameworks belong to the multi-mind framework, where results are the averaged value of 20 groups. The SC-cBCI framework involves signal-level fusion, while the MC-cBCI framework and the MLDANet-cBCI

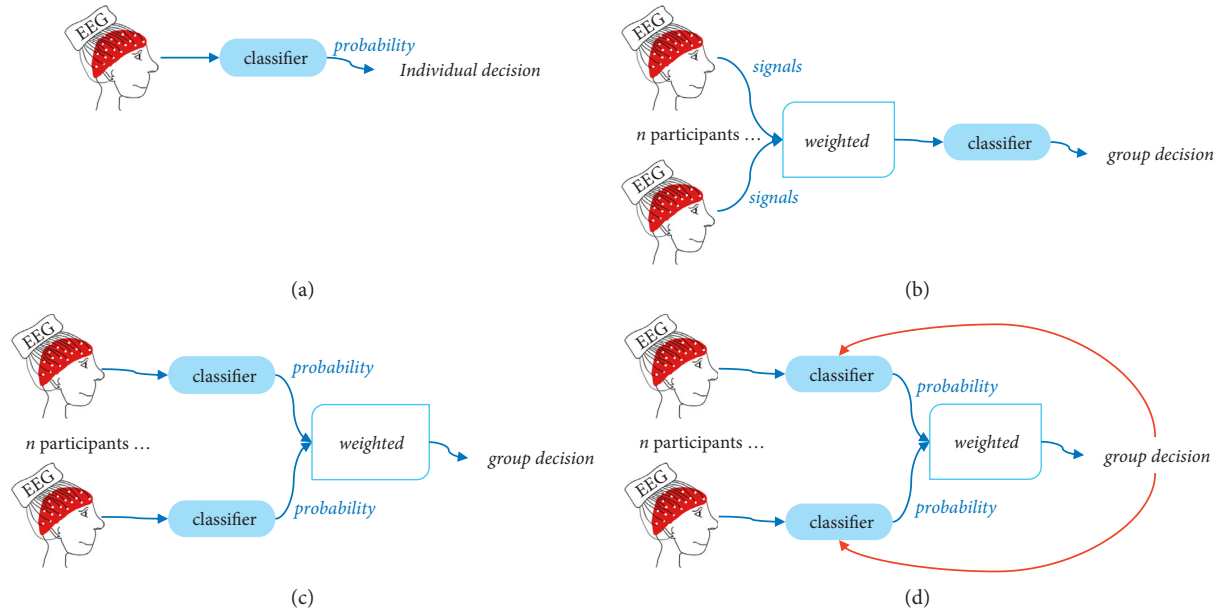


FIGURE 8: Summary of different BCI frameworks: (a) sBCI framework; (b) SC-cBCI framework; (c) MC-cBCI framework; (d) MLDANet-cBCI framework.

framework perform decision-level fusion. In particular, the MLDANet-cBCI framework introduces the mutual learning strategy, and the decision-making is dynamic and interactive.

The optimal parameter values are shown in Table 1, with all frameworks trained on an NVIDIA TITAN RTX GPU in the PyTorch platform. We fit the model using the Adam optimizer with cross-entropy function. The weighted coefficients among the three participants were (1,1,1), where the collaborative decision-making was obtained by averaging three decision-making probabilities.

The detection performances of different BCI frameworks are shown in Table 2, namely, classification accuracy, hit rate, false alarm rate, and F1 score. Here, F1 score is viewed as the main performance evaluation criterion due to the imbalanced classification. The significance level by analysis of variance (ANOVA) was performed between MLDANet-cBCI and other cBCI frameworks. The results indicated that cBCI frameworks outperform the sBCI framework. In cBCI frameworks, decision-level fusion (MC-cBCI and MLDANet-cBCI) outperformed signal-level fusion (SC-cBCI). The MLDANet-cBCI with mutual learning strategy performed the best. Compared with the MC-cBCI framework, the F1 score of MLDANet-cBCI framework improved by 0.12, which highlighted the superiority of the MLDANet-cBCI framework. Relatively, both hit rate and false alarm rate of MLDANet-cBCI framework were lower than those of SC-cBCI framework, which illustrates that MLDANet generates a higher decision threshold. The convergence of training loss and testing F1 score in the MLDANet-cBCI framework from 20 groups is shown in Figure 9. When the iterations exceeded 150 rounds, the training loss was unchanged; when the iterations exceeded 50 rounds, the detection performance became stable.

TABLE 1: Network parameter settings.

Parameters	Value		
	sBCI/SC-cBCI	MC-cBCI	ML-cBCI
P3-sSDA network	1	3	3
Batch size	40	40	40
Learning rate	0.0003	0.0003	0.0003
Epoch	100	100	100
α	0.2	0.2	0.4
γ	0.8	0.8	0.8
β	0.2	0.2	0.2

TABLE 2: Detection performances on different BCI frameworks (** $p < 0.01$).

BCI frameworks	Accuracy	Hit rate	False alarm rate	F1 score
sBCI	0.77	0.63	0.20	0.47(**)
SC-cBCI	0.82	0.80	0.18	0.59(**)
MC-cBCI	0.86	0.69	0.11	0.61(**)
MLDANet-cBCI	0.91	0.72	0.05	0.73

4. Discussion

4.1. Effects of Mutual Learning Strategy on Individual Network Capability. It was expected that the mutual learning strategy would improve the individual network capability of the MLDANet-cBCI framework. Here, the individual detection performance for 20 groups in the MC-cBCI and MLDANet-cBCI frameworks was shown (Table 3). The results indicated that the individual F1 score was 0.66 in the MLDANet-cBCI framework, which is significantly higher than that in the MC-cBCI framework ($p < 0.01$), and could even exceed the group performance in the SC-cBCI framework (0.59) and the MC-cBCI framework (0.61) (Table 2). This finding further confirmed that the proposed mutual learning strategy improves the information interaction and dynamic learning capability of

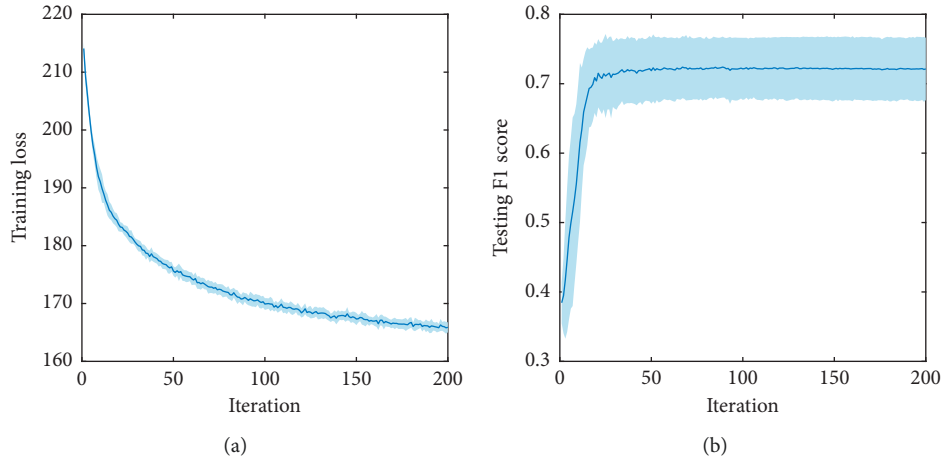


FIGURE 9: The model convergence in the MLDANet-cBCI framework: (a) the convergence of training loss; (b) the convergence of F1 score.

TABLE 3: Detection performance of the individual network in the MC-cBCI and MLDANet-cBCI frameworks.

Groups	Individual F1 score in the MC-cBCI framework			Individual F1 score in the MLDANet-cBCI framework		
	Participant 1	Participant 2	Participant 3	Participant 1	Participant 2	Participant 3
Group 1	0.55	0.37	0.53	0.71	0.63	0.65
Group 2	0.43	0.48	0.37	0.56	0.61	0.53
Group 3	0.42	0.54	0.44	0.33	0.60	0.71
Group 4	0.40	0.49	0.52	0.63	0.69	0.67
Group 5	0.43	0.56	0.45	0.71	0.76	0.76
Group 6	0.45	0.48	0.50	0.70	0.72	0.74
Group 7	0.48	0.41	0.44	0.55	0.55	0.60
Group 8	0.49	0.40	0.50	0.69	0.66	0.68
Group 9	0.38	0.55	0.47	0.59	0.74	0.65
Group 10	0.33	0.53	0.57	0.45	0.78	0.74
Group 11	0.32	0.60	0.52	0.41	0.79	0.70
Group 12	0.50	0.61	0.32	0.70	0.71	0.52
Group 13	0.49	0.41	0.61	0.72	0.73	0.73
Group 14	0.33	0.51	0.50	0.54	0.71	0.70
Group 15	0.49	0.35	0.50	0.69	0.53	0.69
Group 16	0.48	0.33	0.63	0.68	0.61	0.86
Group 17	0.53	0.64	0.41	0.74	0.76	0.62
Group 18	0.51	0.52	0.39	0.73	0.75	0.70
Group 19	0.33	0.60	0.44	0.52	0.75	0.65
Group 20	0.54	0.37	0.42	0.75	0.63	0.62
Average (60 participants)		0.47			0.66	

individual networks in the MLDANet-cBCI framework, where individual networks with so-called poor classification performance can also be cultivated to reach the expert level. This shows that the so-called poor detection performance does not necessarily mean that the data is unreliable; it is likely that the features are not obvious or difficult to extract, and the network training mode also has a great impact on the performance. With the mutual learning strategy, the MLDANet-cBCI framework effectively breaks through the bottleneck of traditional cBCI detection.

4.2. Effects of Number of Individuals in the Source Domain on Detection Performance. In order to achieve a better detection performance, the optimal number of source

domain individuals was evaluated on different BCI frameworks. According to the previous work, source domain individuals with stronger P3 maps contribute to better detection performance [35]. When there were few individuals in the source domain, individuals with stronger P3 maps were preferentially selected as the source domain for better performance. In order to observe the change trend of detection performance with different number of individuals in the source domain and simplify the calculation, source domain individuals were sorted in descending order according to P3 maps of 29 participants. In addition, the first 4, 7, 10, 13, and 16 individuals were used to construct the source domain, as shown in Figure 10. The results indicated that around 13~16 source domain individuals could contribute to the

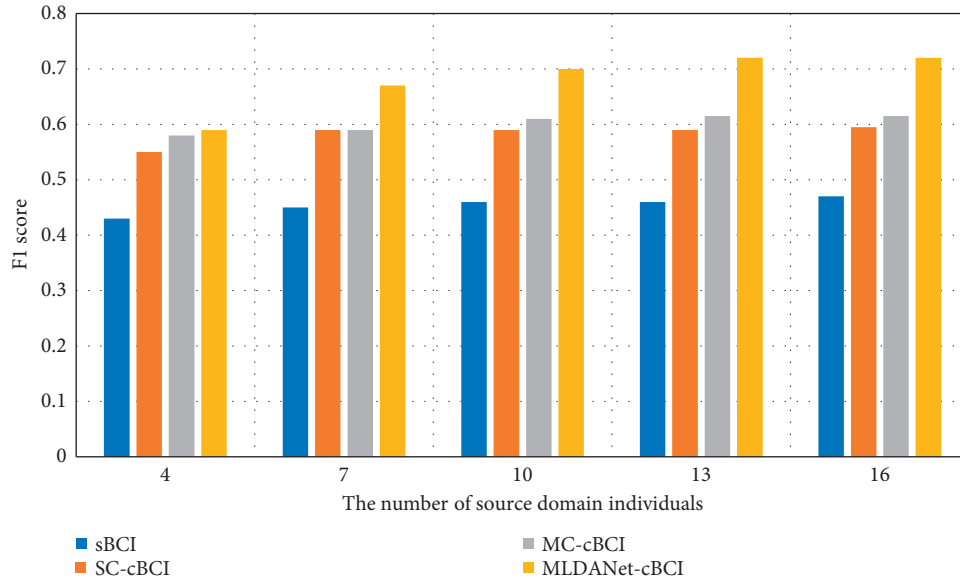


FIGURE 10: Detection performances with different number of individuals in the source domain.

improved performance for the multi-mind EEG signals in this paradigm. Furthermore, the MLDANet-cBCI framework always showed the best performance. Notably, the proposed MLDANet-cBCI framework was particularly sensitive to the number of source domain individuals, where the F1 score improved by 0.14 with their number varying from 4 to 13. Thus, the superiority of MLDANet was based on a sufficient number of source domain individuals.

5. Conclusions

In the present work, we developed a multi-mind cBCI framework to enhance the group detection performance of dynamic visual targets. In this framework, a mutual learning domain adaptation network (MLDANet) was proposed to establish the mechanisms of information interaction and dynamic learning between individual network units and collaborative decision-making. By a mutual learning strategy, the collaborative decision-making could guide the re-decision-making process of the individual network for better and more robust detection performance. The results indicated that the proposed MLDANet-cBCI framework outperforms the other cBCI frameworks with the highest F1 score, 0.73, and classification accuracy, 0.91, when three participants collaborate. The mutual learning strategy can effectively improve the individual network capability. Therefore, the proposed cBCI framework provides a novel multi-mind collaborative mode for the improvement of collaborative work performance, which is of great significance for the progression of research on human augmentation.

Data Availability

The datasets in this study are available on request to the corresponding author.

Ethical Approval

The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Ethical Committee of Henan Provincial People's Hospital.

Consent

Informed consent was obtained from all subjects involved in the study.

Conflicts of Interest

The authors declare no conflicts of interest.

Authors' Contributions

Jun Shu contributed to conceptualization. Xiyu Song participated in methodology, data curation, and visualization. Li Tong helped with validation. Bin Yan was responsible for formal analysis and funding acquisition. Qiang Yang and Jian Kou conducted investigation. Xiyu Song and Ying Zeng prepared and wrote the original draft, and reviewed and edited the manuscript. Ying Zeng was involved in project administration. All authors read and agreed to the published version of the manuscript.

Acknowledgments

This research was funded by the National Key Research and Development Plan of China under grant no. 2017YFB1002502, the National Natural Science Foundation of China (no. 61701089), and the Natural Science Foundation of Henan Province of China (no. 162300410333). The authors would like to thank all the participants who joined the experiment.

References

- [1] Y. Li and C. S. Nam, "Collaborative brain-computer interface for people with motor disabilities [research frontier]," *IEEE Computational Intelligence Magazine*, vol. 11, no. 3, pp. 56–66, 2016.
- [2] J. Pan, Q. Xie, P. Qin et al., "Prognosis for patients with cognitive motor dissociation identified by brain-computer interface," *Brain*, vol. 143, no. 4, pp. 1177–1189, 2020.
- [3] D. Valeriani, C. Cinel, and R. Poli, "Brain-computer interfaces for human augmentation," *Brain Sciences*, vol. 9, no. 22, pp. 1–3, 2019.
- [4] C. Cinel, D. Valeriani, and R. Poli, "Neurotechnologies for human cognitive augmentation: current state of the art and future prospects," *Frontiers in Human Neuroscience*, vol. 13, no. 13, pp. 13–45, 2019.
- [5] J. Jin, Z. Wang, R. Xu, C. Liu, X. Wang, and A. Cichocki, "Robust similarity measurement based on a novel time filter for SSVEPs detection," *IEEE Transactions on Neural Networks and Learning Systems*, pp. 1–10, 2021.
- [6] J. Jin, R. Xiao, I. Daly, Y. Miao, X. Wang, and A. Cichocki, "Internal feature selection method of CSP based on L1-norm and dempster-shafer theory," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 11, pp. 4814–4825, 2021.
- [7] Y. Li, F. Wang, Y. Chen, A. Cichocki, and T. Sejnowski, "The effects of audiovisual inputs on solving the cocktail party problem in the human brain: an fMRI study," *Cerebral Cortex*, vol. 28, no. 10, pp. 3623–3637, 2018.
- [8] 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 Transactions on Biomedical Engineering*, vol. 67, no. 11, pp. 3073–3082, 2020.
- [9] Z. Zhou, E. Yin, Y. Liu, J. Jiang, and D. Hu, "A novel task-oriented optimal design for P300-based brain-computer interfaces," *Journal of Neural Engineering*, vol. 11, no. 5, Article ID 056003, 2014.
- [10] A. J. Solon, S. M. Gordon, J. R. Mcdaniel, and V. J. Lawhern, "Collaborative brain-computer interface for human interest detection in complex and dynamic settings," in *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, pp. 970–975, Miyazaki, Japan, October 2018.
- [11] X. Song, B. Yan, L. Tong, J. Shu, and Y. Zeng, "Asynchronous video target detection based on single-trial EEG signals," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 28, no. 9, pp. 1931–1943, 2020.
- [12] A. Nijholt and M. Poel, "Multi-Brain BCI: characteristics and social interactions," in *Proceedings of the International Conference on Augmented Cognition*, pp. 79–90, Toronto, ON, Canada, July 2016.
- [13] D. Valeriani and A. Matran-Fernandez, "Past and future of multi-mind brain-computer interfaces," *Brain-Computer Interfaces Handbook: Technological and Theoretical Advances*, CRC Press, Boca Raton, FL, USA, 2018.
- [14] G. Dominijanni, S. Shokur, G. Salvietti, S. Buehler, and S. Micera, "Enhancing human bodies with extra robotic arms and fingers: the neural resource allocation problem," 2021, <https://arxiv.org/abs/2103.17252>.
- [15] S. Bhattacharyya, D. Valeriani, C. Cinel, L. Citi, and R. Poli, "Target detection in video feeds with selected dyads and groups assisted by collaborative brain-computer interfaces," in *Proceedings of the 2019 9th International IEEE/EMBS Conference on Neural Engineering (NER)*, pp. 159–162, San Francisco, CA, USA, May 2019.
- [16] D. Valeriani, C. Cinel, and R. Poli, "Group augmentation in realistic visual-search decisions via a hybrid brain-computer interface," *Scientific Reports*, vol. 7, no. 7772, pp. 1–12, 2017.
- [17] R. Poli, C. Cinel, A. Matran-Fernandez, F. Sepulveda, and A. Stoica, "Towards cooperative brain-computer interfaces for space navigation," in *Proceedings of the 2013 International Conference on Intelligent user Interfaces*, pp. 149–160, Santa Monica, CA, USA, March 2013.
- [18] K.-C. Chuang and Y.-P. Lin, "Cost-efficient, portable, and custom multi-subject electroencephalogram recording system," *IEEE Access*, vol. 7, pp. 56760–56769, 2019.
- [19] A. Matran-Fernandez, R. Poli, and C. Cinel, "Collaborative brain-computer interfaces for the automatic classification of images," in *Proceedings of the 6th International IEEE/EMBS Conference on Neural Engineering (NER)*, pp. 1096–1099, San Diego, CA, USA, November 2013.
- [20] A. Matran-Fernandez and R. Poli, "Collaborative brain-computer interfaces for target localisation in rapid serial visual presentation," in *Proceedings of the Computer Science & Electronic Engineering Conference*, pp. 127–132, Denpasar, Indonesia, November 2014.
- [21] A. Stoica, A. Matran-Fernandez, D. Andreou et al., "Multi-brain fusion and applications to intelligence analysis," *Multisensor, Multisource Information Fusion: Architectures, Algorithms, and Applications*, vol. 8756, 2013.
- [22] Y. Wang and T.-P. Jung, "A collaborative brain-computer interface for improving human performance," *PLoS One*, vol. 6, no. 5, Article ID e20422, 2011.
- [23] H. Cecotti and B. Rivet, "Subject combination and electrode selection in cooperative brain-computer interface based on event related potentials," *Brain Sciences*, vol. 4, no. 2, pp. 335–355, 2014.
- [24] H. Cecotti and B. Rivet, "Performance estimation of a cooperative brain-computer interface based on the detection of steady-state visual evoked potentials," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 2059–2063, Florence, Italy, May 2014.
- [25] D. Valeriani, R. Poli, and C. Cinel, "A collaborative brain-computer interface for improving group detection of visual targets in complex natural environments," in *Proceedings of the 7th International IEEE/EMBS Conference on Neural Engineering (NER)*, pp. 25–28, Montpellier, France, April 2015.
- [26] D. Valeriani, R. Poli, and C. Cinel, "Enhancement of group perception via a collaborative brain-computer interface," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 6, pp. 1238–1248, 2016.
- [27] D. Valeriani, C. Cinel, and R. Poli, "Augmenting group performance in target-face recognition via collaborative brain-computer interfaces for surveillance applications," in *Proceedings of the 8th International IEEE/EMBS Conference on Neural Engineering (NER)*, pp. 415–418, Piscataway, NJ, USA, May 2017.
- [28] P. Yuan, Y. Wang, W. Wu, H. Xu, and S. Gao, "Study on an online collaborative BCI to accelerate response to visual targets," in *Proceedings of the 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, vol. 2012, pp. 1736–1739, San Diego, CA, USA, August 2012.
- [29] P. Yuan, Y. Wang, X. Gao, T.-P. Jung, and S. Gao, "A collaborative brain-computer interface for accelerating human decision making," in *Proceedings of the International*

- Conference on Universal Access in Human-Computer Interaction*, pp. 672–681, Las Vegas, NV, USA, July 2013.
- [30] L. Jiang, Y. Wang, B. Cai, W. Chen, and X. Zheng, “Rapid face recognition based on single-trial event-related potential detection over multiple brains,” in *Proceedings of the International IEEE/EMBS Conference on Neural Engineering, NER*, vol. 2015, pp. 106–109, Montpellier, France, April 2015.
 - [31] H. Zhang, L. Zhu, S. Xu, J. Cao, and W. Kong, “Two brains, one target: design of a multi-level information fusion model based on dual-subject RSVP,” *Journal of Neuroscience Methods*, vol. 363, Article ID 109346, 2021.
 - [32] M. P. Eckstein, K. Das, B. T. Pham et al., “Neural decoding of collective wisdom with multi-brain computing,” *NeuroImage*, vol. 59, no. 1, pp. 94–108, 2012.
 - [33] D. Valeriani and R. Poli, “Cyborg groups enhance face recognition in crowded environments,” *PLoS One*, vol. 14, no. 3, Article ID e0212935, 2019.
 - [34] A. Matran-Fernandez and R. Poli, “Towards the automated localisation of targets in rapid image-sifting by collaborative brain-computer interfaces,” *PLoS One*, vol. 12, no. 5, pp. 1–28, 2017.
 - [35] X. Song, Y. Zeng, L. Tong, J. Shu, G. Bao, and B. Yan, “P3-MSDA: multi-source domain adaptation network for the dynamic visual target detection,” *Frontiers in Human Neuroscience*, vol. 15, p. 429, 2021.
 - [36] H. Ramoser, J. Muller-Gerking, and G. Pfurtscheller, “Optimal spatial filtering of single trial EEG during imagined hand movement,” *IEEE Transactions on Rehabilitation Engineering*, vol. 8, no. 4, pp. 441–446, 2000.