Dual-Brain EEG Decoding for Target Detection via Joint Learning in Shared and Private Spaces

Bingfeng He, Li Zhu, Junhua Li, *Senior Member, IEEE*, Andrzej Cichocki *Fellow, IEEE*, and Wanzeng Kong, *Senior Member, IEEE*

***Abstract*—Hyperscanning enables simultaneous electroencephalography (EEG) recording from multiple individuals, facilitating collaborative brain activity to reduce individual biases and enhance the reliability of decision-making. The decoding of such collaborative paradigm tasks has traditionally relied solely on simple fusion methods based on each individual brain activity, without incorporating cross-brain coupling information. Inspired by social interaction studies on enhanced inter-brain synchrony in collaborative tasks using hyperscanning, we propose a joint learning framework for dual-brain target detection that integrates a shared space construction module and shared feature-guided module. The shared space construction module incorporates brain-to-brain coupling analysis to identify cross-brain synchrony, and further integrates shared and private features through a multi-head fusion mechanism for joint representation learning in shared feature-guided module. Experimental results show an average 10% improvement in balanced accuracy across 12 participant groups compared to traditional single-brain approaches, with some groups achieving up to a 5% gain over state-of-the-art (SOTA) methods. Notably, higher-performing groups exhibit stronger inter-brain coupling and more synchronized target-related responses. These findings advance the development of collaborative brain-computer interface (BCI) systems for more robust and effective target detection.**

***Index Terms*—EEG-based hyperscanning, cross-brain synchrony, shared space, multi-head attention**

I. Introduction

BRAIN-COMPUTER Interface (BCI) systems enable direct communication between computers and the human brains without relying on peripheral nerves or muscles [1]. Among the neurophysiological signals used for BCI, Electroencephalography (EEG)-based systems have gained significant attention due to their noninvasive nature, ease of operation, and high temporal resolution [2]. One of the key paradigms in EEG-based BCI is Rapid Serial Visual Presentation (RSVP), which has been widely applied in areas such as image retrieval, anti-spoofing, and anomaly detection [3], [4]. Unlike traditional image retrieval methods that depend solely on computer vision and often struggle with complex or camouflaged targets, RSVP-based BCI systems mitigate these challenges by utilizing EEG-derived neural responses.

In RSVP paradigm, sequential image streams containing occasional target stimuli elicit specific event related potential (ERP) components such as P300, N200, and late positive potential (LPP), which reflect cognitive processes related to attention and stimulus evaluation [5], [6], [7], [8]. However, effective ERP extraction is hindered by low signal-to-noise ratios, physiological artifacts (e.g., eye/muscle activities), and temporal overlap of stimuli, along with individual variability in attention and fatigue levels [9], [10]. These limitations constrain the robustness and scalability of RSVP-BCIs, particularly in real-world, dynamic environments.

Collaborative BCI decoding: Inter-brain dynamics have shown potential to augment system reliability and performance by leveraging group-level neural correlations through hyperscanning[11], [12]. To address the limitations of individual variability and improve the robustness of neural signal decoding, recent studies have explored collaborative BCI (cBCI) paradigms using hyperscanning, which allows for the simultaneous acquisition of EEG data from multiple individuals engaged in a cooperative task [13], [14]. This paradigm introduces an additional layer of inter-brain information, enabling the investigation of collective cognitive mechanisms and temporally synchronized neural responses [15]. Despite these promising directions, most existing decoding frameworks primarily focus on fusing features or decisions at the individual level within a group [16], [17], [18], [19], [20]. These methods often ignore informative cross-brain coupling features, which are critical for capturing shared representations and neural synchrony during cooperative tasks. Recent work by Falcon-Caro *et al.* [21] highlights the importance of constructing joint spatial filters across multiple users to capture shared neural structures while preserving subject-specific characteristics—underscoring the need for more principled decoding architectures in multi-brain BCI. In parallel, insights from social neuroscience have demonstrated that inter-brain synchrony is significantly enhanced during collaborative interactions compared to competitive or isolated conditions [22], [23], [24], [25]. These findings suggest that collaborative contexts naturally elicit stronger cross-brain connectivity, which can be leveraged to improve decoding performance in cBCI systems.

Our work: Building upon the growing recognition of shared information as a critical component in collaborative

This work was supported the National Science Foundation of China under Grant Nos. 62471169, 62301196; in part by the "Leading Goose" R&D Program of Zhejiang under Grant No.2023C03026 and Zhejiang Provincial Natural Science Foundation of China under Grant No.LO24F020035. (*Corresponding author: Wanzeng Kong.)*

Approval of all ethical and experimental procedures and protocols was granted by the Research Ethics Committee of Saitama Institute of Technology under Approval No.2018-01.

Bingfeng He, Li Zhu and Wanzeng Kong are with the School of Computer Science and Technology, Hangzhou Dianzi University, Hangzhou 310018, China (email: kongwanzeng@hdu.edu.cn).

Junhua Li is with the School of Computer Science and Electronic Engineering, University of Essex, CO4 3SQ Colchester, U.K., and also with the Laboratory for Brain-Bionic Intelligence and Computational Neuroscience, Wuyi University, Jiangmen 529020, China.

Andrzej Cichocki is with Systems Research Institute, Nicolaus Copernicus University, Poland and RIKEN (AIP), Japan.

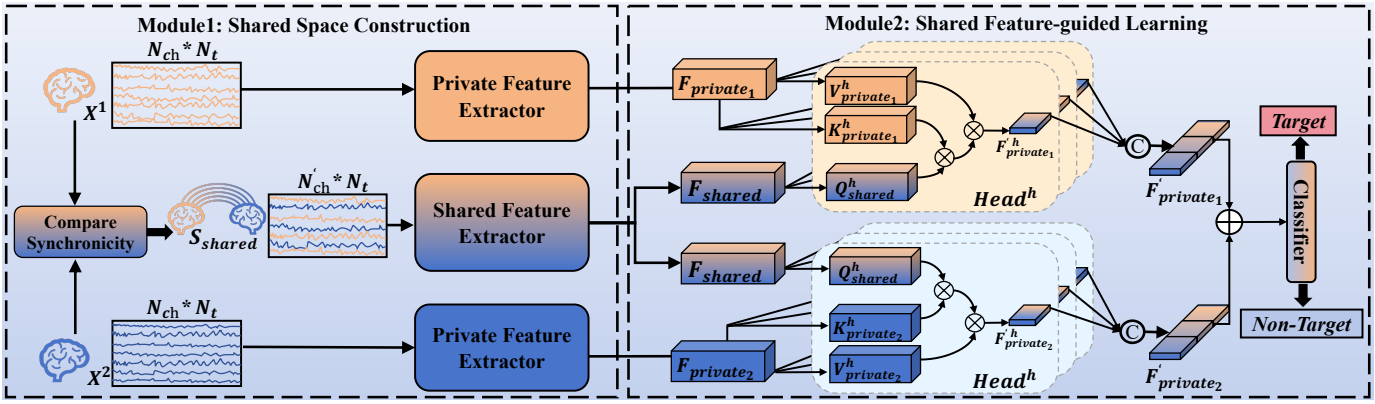

Fig. 1: Overview of the proposed dual-brain EEG decoding for target detection framework. Module 1 constructs a shared space via synchronized EEG analysis and extracts shared and private features respectively. Module 2 applies multi-head attention using shared feature to guide the fusion of private features.

BCI paradigms, this paper presents a novel joint learning framework for EEG-based hyperscanning target detection. The proposed approach addresses the limitations of conventional fusion-based methods by incorporating both shared and private representations for enhanced decoding. (1) We develop a joint learning framework that constructs a shared representational space through coupling analysis along with task-related features, which complements private decoding by providing group-level contextual information. (2) A dedicated hyperscanning experiment was conducted to generate a dual-subject EEG dataset, enabling comprehensive evaluation of the proposed method. (3) We propose a shared feature-guided learning strategy that injects shared representations into the private decoding pathway during training, enhancing subject-specific feature extraction and improving overall model robustness.

II. Method

We propose a hyperscanning-based framework for collaborative dual-brain target detection, based on EEG signals recorded from paired subjects. The method consists of two main modules. Module 1 captures inter-subject neural synchronization to extract both private and shared representations. Module 2 enhances decoding by integrating shared features into a multi-head attention mechanism, enabling more effective fusion of individual representations for classification. An overview is shown in Fig. 1.

A. Module 1: Shared Space Construction

EEG data are collected via hyperscanning to capture mutual cognitive processes between two participants, ensuring temporal synchronization. The pre-processed data are denoted as $\boldsymbol{X}^p \in \mathbb{R}^{N_{ch} \times N_t}$ $(p \in \{1, 2\})$, where N_{ch} is the number of EEG channels and N_t is the number of time points. For each participant p, the time series from the i-th channel is denoted as $x_i^p = \boldsymbol{X}^p[i, :] \in \mathbb{R}^{N_t}$, where $i = 1, \ldots, N_{ch}$.

To quantify shared neural information, phase synchronization is computed using the phase-locking value (PLV):

$$PLV_i = \frac{1}{N_t} \left| \sum_{t=1}^{N_t} e^{j(\phi^1[i,t] - \phi^2[i,t])} \right|, \tag{1}$$

where $\phi^1[i, t]$ and $\phi^2[i, t]$ are the instantaneous phases of the i-th channel for each participant, which are computed using the Hilbert transform applied to the EEG signals. The resulting PLV values, which quantify phase alignment, are arranged into a vector $\boldsymbol{S}$. High synchronization values are identified by sorting $\boldsymbol{S}$ in descending order and selecting the top n indices, $\{i_1, i_2, \ldots, i_n\}$. For each selected channels, the corresponding time-series data are concatenated to form the shared space representation:

$$\mathbf{z}_{ik} = \begin{bmatrix} x_{ik}^1 & x_{ik}^2 \end{bmatrix}^\top, \quad k = 1, 2, \ldots, n, \tag{2}$$

$$\mathbf{S}_{\text{shared}} = \begin{bmatrix} \mathbf{z}_{i1}^\top & \cdots & \mathbf{z}_{in}^\top \end{bmatrix}^\top \in \mathbb{R}^{N'_{ch} \times N_t}. \tag{3}$$

This shared space captures synchronized neural activity and serves as input to a multi-scale convolutional network. The network processes three inputs: $\boldsymbol{X}^1$, $\boldsymbol{X}^2$, and $\boldsymbol{S}_{shared}$, each through an independent feature extractor employing multi-scale convolutional blocks to extract both temporal and spatial neural features, as illustrated in the multi-scale feature extraction block in Fig. 2.

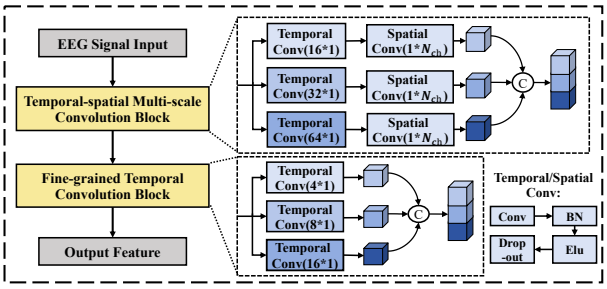

Fig. 2: The multi-scale convolutional feature extractor leverages temporal and spatial convolutions to capture target detection task-related neural features. Each convolutional block is followed by batch normalization (BN), exponential linear activation (ELU), and dropout regularization.

The feature extractor receives inputs from both the shared and private spaces, and independently extracts feature vectors from each. This multi-scale convolutional architecture effectively captures both temporal and spatial features. The first

convolutional block extracts broad temporal patterns, while the second refines these patterns to capture more granular details. The tailored depthwise convolution kernel sizes ensure compatibility with the input dimensions, and the use of average pooling minimizes computational overhead while retaining crucial information, providing a robust foundation for subsequent tasks such as collaborative decision-making and target classification.

B. Module 2: Shared Feature-guided Learning

Building upon the comprehensive neural features extracted in Module 1 using the multi-scale convolutional network, this module focuses on leveraging the shared feature representation $\boldsymbol{F}_{shared}$ to enhance the private features of each participant, denoted as $\boldsymbol{F}_{private_1}$ and $\boldsymbol{F}_{private_2}$. The objective is to integrate relevant shared information into the individual representations, thereby improving the model's ability to capture task-relevant patterns embedded in EEG signals.

To achieve this, a multi-head attention mechanism is employed. It projects both shared and private feature vectors into query, key, and value spaces across multiple attention heads, allowing the model to compute attention scores to quantify the influence of shared features on each participant's private representation. For the h-th attention head, the shared feature vector is linearly transformed into a query matrix:

$$\boldsymbol{Q}^h_{shared} = \boldsymbol{F}_{shared} \times \boldsymbol{W}^h_q, \tag{1}$$

where $\boldsymbol{W}^h_q$ is a learnable weight matrix for query transformation.

The private feature vectors $\boldsymbol{F}_{private_p}(p \in \{1,2\})$ are linearly projected into key and value matrices for attention computation:

$$\boldsymbol{K}^h_{private_p} = \boldsymbol{F}_{private_p}\boldsymbol{W}^h_k, \tag{2}$$
$$\boldsymbol{V}^h_{private_p} = \boldsymbol{F}_{private_p}\boldsymbol{W}^h_v. \tag{3}$$

The attention scores are computed using the scaled dot-product mechanism across multiple heads:

$$\boldsymbol{H}^h_{private_p} = \text{Softmax}\left(\frac{\boldsymbol{Q}^h_{shared}\boldsymbol{K}^{h\top}_{private_p}}{\sqrt{d_k}}\right)\boldsymbol{V}^h_{private_p}, \tag{4}$$

where d_k is the dimensionality of the key matrix, used to normalize the dot product for numerical stability.

The outputs from all heads are concatenated to form the enhanced private features:

$$\boldsymbol{F}^{'}_{private_p} = \text{Concat}(\boldsymbol{H}^1_{private_p}, \ldots, \boldsymbol{H}^h_{private_p}), \tag{5}$$

where h represents the number of attention heads, and $\boldsymbol{F}^{'}_{private_p}$ denotes the enhanced version of $\boldsymbol{F}_{private_p}$ guided by the shared features.Finally, the enhanced private features from both participants are concatenated and passed through a fully connected layer for target classification.

The proposed multi-head attention mechanism selectively incorporates shared features to enhance individual representations without causing interference. This design effectively balances individual-specific and collaborative neural patterns, thereby improving the robustness and task relevance of EEG decoding in dual-brain settings.

III. Experiments

A. Experimental Design

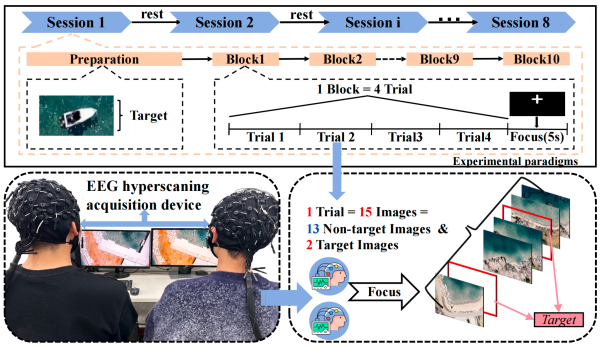

Fig. 3: Illustration of the experimental paradigm and scenario.

The objective of this experimental design is to integrate a multi-brain framework with a classical RSVP paradigm. For the multi-brain component, EEG signals from two participants were simultaneously recorded using a hyperscanning setup. In the RSVP task, 24 right-handed participants were randomly assigned into 12 pairs. The experiment consisted of eight sessions. As illustrated in Fig. 3, during each session, participants were instructed to search for a white yacht within a simulated panoramic satellite image presented from a top-down aerial perspective. Each image was displayed for 800 ms.

EEG signals were recorded using two 64-channel Neuroscan systems following the international 10–20 system. Signals were referenced to the vertex, grounded at the forehead, and maintained below 15 kΩ impedance. Data were initially sampled at 1000 Hz, then band-pass filtered between 0.1 and 40 Hz, and subsequently downsampled to 250 Hz. Bad channels were interpolated using the MNE library [26] to minimize data loss and ensure data quality.

B. Comparison of Decoding Performance Between Dual-Brain and Single-Brain Modes

We compare the decoding performance of the dual-brain RSVP paradigm with that of the single-brain paradigm. In the target detection task, the ratio of target to non-target images is 13:2, resulting in a data imbalance. To address this, we adopt balanced accuracy (BA) as the evaluation metric. In the single-brain setting, BA is computed as the arithmetic mean of the two participants' individual scores. As shown in Fig. 4, the dual-brain setup outperformed the single-brain setup by an average of over 10.00% in BA. Moreover, the dual-brain paradigm achieved superior performance across all participant pairs, with the best-performing group showing a 14.17% higher BA compared to the single-brain mode. These results highlight the effectiveness of using shared-space information to guide private-space representation integration.

TABLE I: Performance Comparison of Different Methods across Groups (BA/%)

Method	Subjects Groups												Average
	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11	G12	
LMDA-net	80.94	90.79	85.23	78.23	77.52	88.14	82.41	75.62	78.94	77.67	89.51	81.34	82.19
EEGNet	82.89	91.68	89.21	81.25	79.29	91.40	87.16	80.20	82.43	83.42	91.42	82.40	85.23
EEGInception	86.95	94.92	92.65	84.87	82.51	93.58	89.05	83.37	83.37	86.88	95.44	85.44	88.25
HyperscanNet	83.45	91.98	90.10	82.41	80.21	90.65	89.12	82.68	82.46	83.43	92.41	82.15	85.02
PLNet	82.79	92.15	87.74	82.45	80.55	92.14	86.24	81.54	81.25	85.12	92.15	81.90	85.51
EEG-Conformer	86.91	96.03	88.43	86.39	84.13	95.42	90.12	77.45	82.55	82.51	88.32	83.31	86.80
Ours	**89.26**	**97.12**	**94.21**	**86.36**	**85.28**	**96.67**	**91.69**	**84.56**	**85.53**	**90.44**	**97.49**	**87.51**	**90.51**

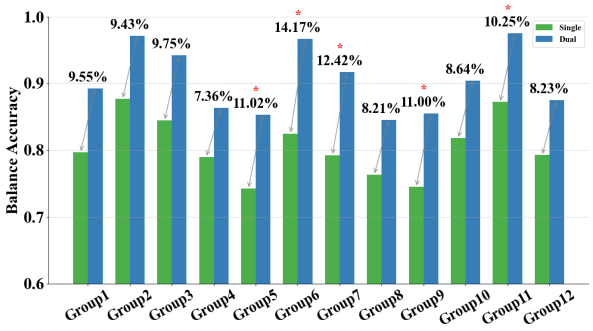

Fig. 4: BA comparison between single-brain and dual-brain configurations, demonstrating significant improvements in target detection with dual-brain collaboration.

TABLE II: Results of Ablation Experiment (%)

Methods	F1-score	BA
Without Module1	80.46 ± 7.24	82.51 ± 6.92
Without Module2	84.91 ± 5.36	85.94 ± 5.62
Full (Ours)	**88.14 ± 3.68**	**90.51 ± 4.48**

C. Comparison with State-of-the-Art Methods

We compared the proposed EEG decoding method with several state-of-the-art (SOTA) approaches, including methods such as LMDA-Net [27], EEGNet [28], EEGInception [29], HyperscanNet [30], PLNet [31], and EEG-Conformer [32].

Table I presents the dual-brain RSVP decoding results across 12 subject pairs. EEGNet, HyperscanNet, and PLNet achieved balanced accuracies (BA) of 85.23%, 85.02%, and 85.51%, respectively, while EEGInception reached 88.25%, and EEG-Conformer 86.80%. In contrast, our method achieved the highest average BA of 90.51%, with notable gains of approximately 5% over PLNet and HyperscanNet.

The advantage is clear in Groups 1, 2, 3, and 11, where our method consistently outperforms all baselines, demonstrating the effectiveness of shared neural representation integration in hyperscanning EEG decoding.

D. Ablation Experiment

This section evaluates the contribution of each module through ablation experiments by selectively removing subcomponents. As shown in Table II, the full model outperforms all variants, confirming the effectiveness of joint learning from shared and private spaces.

Removing Module 1 leads to a performance drop of 8.00% in BA (90.51% vs. 82.51%) and 7.68% in F1-score (88.14% vs. 80.46%), while removing Module 2 results in a 4.57% and 3.23% decrease, respectively. These results demonstrate that both modules substantially enhance classification performance.

E. The synchronization measurement analysis

To validate the relationship between decoding performance and event-related potentials (ERPs), we conducted an ERP analysis to assess inter-subject task synchronization. As shown in Fig. 5, the high-performance group (Group 2) exhibited stronger cross-brain synchronization in the P300 component during target detection compared to the low-performance group (Group 5). As a supplementary analysis to support our main results, these findings not only confirm that the dual-brain EEG data contain task-relevant ERP responses, but also, for the first time, provide empirical evidence that cross-brain neural synchronization is aligned with decoding performance in a collaborative BCI setting.

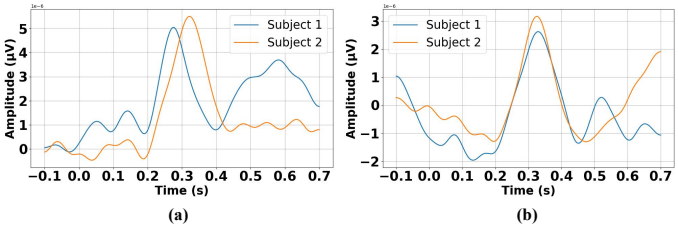

Fig. 5: Comparison of task-related ERP responses between two typical groups: (a) Group 5 , (b) Group 2.

IV. Conclusion

This paper presents an EEG-based hyperscanning dual-brain target detection framework that integrates shared and private space learning for joint decision-making. Experimental results confirm the superiority of the dual-brain paradigm over the single-brain mode, emphasizing the role of shared-space feature-guided private representation learning in enhancing RSVP classification performance.

The core concepts of our framework can be extended to multi-subject (more than 2) scenarios. In particular, the shared space construction can be expanded by incorporating phase synchronization analyses across multiple brains, while the multi-head attention mechanism could be modified to accommodate additional subjects by enhancing the fusion of features from multiple participants.

REFERENCES

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

[2] R. Janapati, V. Dalal, and R. Sengupta, "Advances in modern eeg-bci signal processing: A review," *Materials Today: Proceedings*, vol. 80, pp. 2563–2566, 2023.

[3] J. Mao, S. Qiu, W. Wei, and H. He, "Cross-modal guiding and reweighting network for multi-modal rsvp-based target detection," *Neural Networks*, vol. 161, pp. 65–82, 2023.

[4] H. Wu, F. Li, W. Chu, Y. Li, Y. Niu, G. Shi, L. Zhang, and Y. Chen, "Semantic image sorting method for rsvp presentation," *Journal of Neural Engineering*, vol. 21, no. 3, p. 036018, 2024.

[5] Q. Zhou, Q. Zhang, B. Wang, Y. Yang, Z. Yuan, S. Li, Y. Zhao, Y. Zhu, Z. Gao, J. Zhou *et al.*, "Rsvp-based bci for inconspicuous targets: detection, localization, and modulation of attention," *Journal of Neural Engineering*, vol. 21, no. 4, p. 046046, 2024.

[6] W. Wei, S. Qiu, Y. Zhang, J. Mao, and H. He, "Erp prototypical matching net: A meta-learning method for zero-calibration rsvp-based image retrieval," *Journal of Neural Engineering*, vol. 19, no. 2, p. 026028, 2022.

[7] S. Park, J. Ha, and L. Kim, "Event-related pupillary response-based authentication system using eye-tracker add-on augmented reality glasses for individual identification," *Frontiers in Physiology*, vol. 15, p. 1325784, 2024.

[8] Á. Fernández-Rodríguez, C. Álvarez, C. Langin, F. Velasco-Álvarez, T. Letouzé, J.-M. André, and R. Ron-Angevin, "Exploring the impact of stimulus transparency in erp-bci under rsvp," in *Proceedings of the 17th International Conference on PErvasive Technologies Related to Assistive Environments*, 2024, pp. 9–14.

[9] Y. Cui, S. Xie, X. Xie, D. Zheng, H. Tang, K. Duan, X. Chen, and Y. Jiang, "Lder: a classification framework based on erp enhancement in rsvp task," *Journal of Neural Engineering*, vol. 20, no. 3, p. 036029, 2023.

[10] G. Li, S. Huang, W. Xu, W. Jiao, Y. Jiang, Z. Gao, and J. Zhang, "The impact of mental fatigue on brain activity: A comparative study both in resting state and task state using eeg," *BMC neuroscience*, vol. 21, pp. 1–9, 2020.

[11] F. Babiloni and L. Astolfi, "Social neuroscience and hyperscanning techniques: past, present and future," *Neuroscience & Biobehavioral Reviews*, vol. 44, pp. 76–93, 2014.

[12] N. Sciaraffa, J. Liu, P. Aricò, G. D. Flumeri, B. M. Inguscio, G. Borghini, and F. Babiloni, "Multivariate model for cooperation: bridging social physiological compliance and hyperscanning," *Social Cognitive and Affective Neuroscience*, vol. 16, no. 1-2, pp. 193–209, 2021.

[13] L. Zhu, Y. Liu, R. Liu, Y. Peng, J. Cao, J. Li, and W. Kong, "Decoding multi-brain motor imagery from eeg using coupling feature extraction and few-shot learning," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 31, pp. 4683–4692, 2023.

[14] C.-H. Chuang, P.-H. Peng, and Y.-C. Chen, "Leveraging hyperscanning eeg and vr omnidirectional treadmill to explore inter-brain synchrony in collaborative spatial navigation," *arXiv preprint arXiv:2406.06327*, 2024.

[15] T. Jia, J. Sun, C. McGeady, L. Ji, and C. Li, "Enhancing brain–computer interface performance by incorporating brain-to-brain coupling," *Cyborg and Bionic Systems*, vol. 5, p. 0116, 2024.

[16] S. Cheng and J. Wang, "Mi 2 mi: Training dyad with collaborative brain-computer interface and cooperative motor imagery tasks for better bci performance," *arXiv preprint arXiv:2406.00470*, 2024.

[17] Z. Zhao, Y. Lin, Y. Wang, and X. Gao, "Single-trial eeg classification using spatio-temporal weighting and correlation analysis for rsvp-based collaborative brain computer interface," *IEEE Transactions on Biomedical Engineering*, vol. 71, no. 2, pp. 553–562, 2023.

[18] P. Li, J. Su, A. N. Belkacem, L. Cheng, and C. Chen, "Corrigendum: Multi-person feature fusion transfer learning-based convolutional neural network for ssvep-based collaborative bci," *Frontiers in Neuroscience*, vol. 16, p. 1024150, 2022.

[19] ——, "Steady-state visually evoked potential collaborative bci system deep learning classification algorithm based on multi-person feature fusion transfer learning-based convolutional neural network." *Frontiers in Neuroscience*, vol. 16, 2022.

[20] S. Rao, D. C. Raju, P. N. Sai, M. Prabhakar, A. Tiwari, J. S. Dhatterwal, and U. K. Shukla, "Real-time collaborative gaming using multiagent brain-computer interfaces," in *2023 International Conference on Communication, Security and Artificial Intelligence (ICCSAI)*. IEEE, 2023, pp. 705–709.

[21] A. Falcon-Caro, S. Shirani, J. F. Ferreira, J. J. Bird, and S. Sanei, "Formulation of common spatial patterns for multi-task hyperscanning bci," *IEEE Transactions on Biomedical Engineering*, vol. 71, no. 6, pp. 1950–1957, 2024.

[22] G. A. M. Vasiljevic and L. C. de Miranda, "Comparing users' performance and game experience between a competitive and collaborative brain-computer interface," *Behaviour & Information Technology*, vol. 43, no. 1, pp. 40–59, 2024.

[23] S. Zhou, Y. Zhang, Y. Fu, L. Wu, X. Li, N. Zhu, D. Li, and M. Zhang, "The effect of task performance and partnership on interpersonal brain synchrony during cooperation," *Brain Sciences*, vol. 12, no. 5, p. 635, 2022.

[24] Q. Liu, H. Cui, B. Huang, Y. Huang, H. Sun, X. Ru, M. Zhang, and W. Chen, "Inter-brain neural mechanism and influencing factors underlying different cooperative behaviors: a hyperscanning study," *Brain Structure and Function*, vol. 229, no. 1, pp. 75–95, 2024.

[25] K. A. Usal and M. P. Çakır, "Modulation of interbrain and intrabrain connectivity due to social presence and task difficulty: A dual eeg/fnirs hyperscanning study," *European Journal of Neuroscience*, vol. 61, no. 7, p. e70091, 2025.

[26] A. Gramfort, M. Luessi, E. Larson, D. A. Engemann, D. Strohmeier, C. Brodbeck, L. Parkkonen, and M. S. Hämäläinen, "Mne software for processing meg and eeg data," *neuroimage*, vol. 86, pp. 446–460, 2014.

[27] Z. Miao, M. Zhao, X. Zhang, and D. Ming, "Lmda-net: A lightweight multi-dimensional attention network for general eeg-based brain-computer interfaces and interpretability," *NeuroImage*, vol. 276, p. 120209, 2023.

[28] V. J. Lawhern, A. J. Solon, N. R. Waytowich, S. M. Gordon, C. P. Hung, and B. J. Lance, "Eegnet: a compact convolutional neural network for eeg-based brain–computer interfaces," *Journal of neural engineering*, vol. 15, no. 5, p. 056013, 2018.

[29] E. Santamaria-Vazquez, V. Martinez-Cagigal, F. Vaquerizo-Villar, and R. Hornero, "Eeg-inception: a novel deep convolutional neural network for assistive erp-based brain-computer interfaces," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 28, no. 12, pp. 2773–2782, 2020.

[30] 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, p. 109346, 2021.

[31] B. Zang, Y. Lin, Z. Liu, and X. Gao, "A deep learning method for single-trial eeg classification in rsvp task based on spatiotemporal features of erps," *Journal of Neural Engineering*, vol. 18, no. 4, p. 0460c8, 2021.

[32] Y. Song, Q. Zheng, B. Liu, and X. Gao, "Eeg conformer: Convolutional transformer for eeg decoding and visualization," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 31, pp. 710–719, 2022.