2025 International Conference on Intelligent Computing
July 26-29, Ningbo, China
https://www.ic-icc.cn/2025/index.php

DASTCN: Enhancing Cross-Subject P300 Detection via Adversarial Spatio-Temporal Learning and Adaptive Source Selection

Xiaodong Yang[1], Fei Wang[2(✉)], and Zhibin Du[3(✉)]

[1] South China Normal University, GuangZhou 528225, China
[2] South China Normal University, GuangZhou 528225, China
fwang@scnu.edu.cn
[3] South China Normal University, GuangZhou 528225, China
zhibindu@126.com

Abstract. Brain-Computer Interface (BCI) systems aim to decode neural activity and translate it into actionable commands for external devices. Electroencephalogram (EEG) is a widely used, non-invasive method for analyzing brain activity. However, the significant inter-subject variability in EEG signals poses a major challenge for the generalization of EEG-based models. While Domain-Adversarial Neural Networks (DANN) have demonstrated promising results in transfer learning tasks, their application to EEG-based cross-subject P300 detection remains relatively unexplored. In this study, we introduce the Domain-Adversarial Spatio-Temporal Convolution Network (DASTCN), which combines a Generative Adversarial Network (GAN) with a lightweight spatio-temporal convolutional architecture to address the issue of inter-subject variability. Extensive empirical evaluations show that DASTCN outperforms conventional models, achieving an accuracy of 84.9% in cross-subject P300 detection. These findings underscore the potential of DASTCN as a transformative tool for advancing practical BCI systems and offer significant implications for future research and applications in this field.

Keywords: Brain-Computer Interface, DANN, GAN, Cross-Subject.

1 Introduction

BCI systems have emerged as a transformative technology, enabling direct communication between the human brain and external devices without the need for peripheral nervous system involvement [1]. By decoding neural signals, BCI technology translates brain activity into actionable commands, offering significant potential in assistive technologies and neurorehabilitation. Among the various neuroimaging modalities, Electroencephalogram (EEG)-based BCI have gained prominence due to their non-invasive nature, portability, and real-time processing capabilities. Prominent EEG-based paradigms include P300, steady-state visual evoked potentials (SSVEP), and motor imagery

(MI) systems. This study focuses on P300-based BCI, which have demonstrated robust performance across diverse applications [2].

A critical challenge in EEG-based BCI is the inherent variability in neural responses across individuals. This inter-subject variability is particularly pronounced in the P300 component, which exhibits significant differences in amplitude, latency, and spatial distribution [3]. Traditional BCI systems typically require extensive user-specific calibration, wherein new users must provide substantial training data to tailor the model to their unique neural patterns. However, this calibration process is often time-consuming and impractical for real-world deployment. Consequently, reducing the reliance on user-specific calibration and developing robust cross-subject adaptation methods has become a pivotal research direction in advancing BCI technology [4].

Significant efforts have been made to minimize or eliminate the calibration phase. For instance, semi-supervised learning approaches, such as those proposed by [5], leverage small labeled datasets to construct support vector machine (SVM) classifiers. Similarly, unsupervised online methods have been applied to develop zero-training BCI for both motor imagery and P300 paradigms [6]. While these methods reduce calibration time, they often require an adaptation period during which system performance gradually improves. Recent advancements in transfer learning, such as the template-based approach introduced by [7], have shown promise in transferring SSVEP templates from existing subjects to new users, thereby enhancing detection accuracy. Additionally, proposed a multimodal EEG-EOG analysis framework for cognitive load assessment [8], significantly improving EEG-based drowsiness detection by mitigating inter- and intra-subject variability.

From the perspective of feature extraction and classifier design, researchers have explored various techniques to enhance BCI performance. For example, utilized spectral feature analysis for cognitive workload classification [9], achieving robust performance in EEG-based BCI. Similarly, applied Common Spatial Pattern (CSP) for feature extraction and used extreme learning machines (ELM) to identify mental workload shifts [10]. Hybrid approaches, such as the integration of EEG with forehead electrooculography (EOG) signals, have also been investigated to evaluate cognitive load levels [11]. Despite their effectiveness in subject-specific contexts, these methods face significant challenges in cross-subject scenarios, primarily due to the high variability in EEG signals across individuals.

The impressive success of deep learning across a wide range of machine learning tasks has spurred its adoption in brain signal analysis. Early work by demonstrated the potential of convolutional neural networks (CNN) for P300 detection [12], employing a four-layer architecture to extract spatial and temporal features from EEG signals. However, the performance of CNN is heavily reliant on the quantity and quality of training data, which is often limited in P300 tasks due to the high cost and difficulty of data collection [13]. This issue is further compounded in BCI research, where the scarcity of large, high-quality datasets remains a significant barrier. To overcome these challenges, proposed EEGNet [14], a generalized deep network utilizing depthwise separable convolutions, which achieved state-of-the-art performance across various EEG detection tasks. Recent studies have further validated the effectiveness of lightweight CNN architectures in cross-subject EEG analysis [15][16].

Despite these advancements, developing a subject-independent BCI system that maintains consistently high performance remains a major challenge. The inherent subject-specific variability, dynamic nature, and weak signal characteristics of EEG, coupled with a low signal-to-noise ratio (SNR), make robust analysis difficult. Moreover, the majority of previous studies have not fully explored the potential of transfer learning techniques. DANN, which have demonstrated significant success in natural language processing and image classification [17], present a promising approach for cross-domain adaptation in EEG analysis. However, applying DANN to cross-subject BCI analysis introduces unique challenges, including class imbalance between source and target domains and the risk of "negative transfer" due to significant inter-subject variability [18].

To address these limitations, we propose the Domain-Adversarial Spatio-Temporal Convolution Network (DASTCN), which introduces three key innovations: (1) A streamlined spatiotemporal feature extraction network for efficient capture of spatial and temporal patterns in P300 signals; (2) A generative adversarial mechanism that synthesizes target-like data from random noise distributions, effectively bridging the source-target domain gap; (3) An adaptive source selection strategy that identifies optimal source subjects through distribution similarity metrics, enhancing cross-subject generalization while mitigating negative transfer effects.

2 Materials

2.1 Subjects

Twenty healthy participants (9 males, and 11 females, aged 18 to 32 years) were recruited for this experiment. The study received approval from the relevant Ethics Committee, and all participants provided written informed consent. During the experiment, participants were instructed to focus on the flashing of a designated target character and mentally count the flashes while their EEG signals were recorded. Each participant completed 30 trials, and the sequence of 30 target characters was randomly generated for each participant before the experiment.

2.2 Experimental paradigm

The electroencephalogram (EEG) signals were acquired at a sampling frequency of 250 Hz utilizing a 30-electrode cap arranged according to the extended 10-20 international system. The acquired signals were referenced to the right mastoid. Signal amplification was performed using a 64-channel SynAmps2 amplifier (Compumedics, Neuroscan, Australia). Throughout the experimental procedure, electrode impedances were meticulously maintained below 5 kΩ, with continuous real-time monitoring facilitated through a dedicated computer interface.

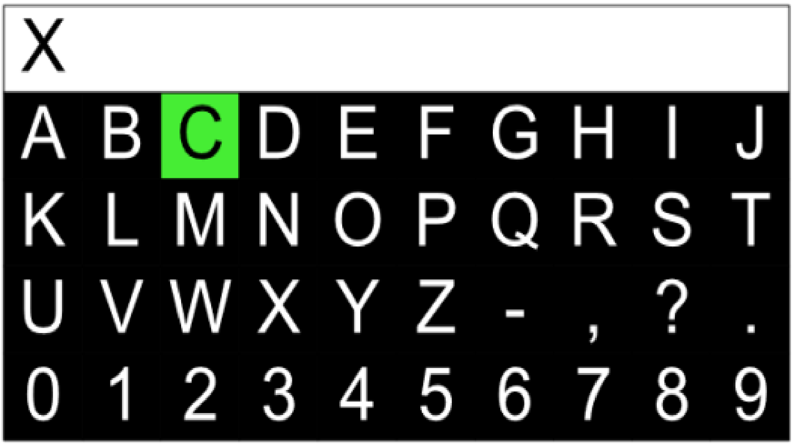

Fig. 1. The GUI of the used P300 speller BCI.

The experimental paradigm employed a P300 speller interface, the graphical representation of which is illustrated in **Fig. 1** [19]. The interface comprised a 4 × 10 matrix containing alphanumeric characters, with participants instructed to focus their attention on specific target characters. Each experimental trial commenced with a 3-second preparatory phase, during which the interface remained static without any visual intensification. Following this preparatory interval, the stimulus presentation phase was initiated, characterized by the sequential illumination of all 40 buttons in a randomized order. This randomization protocol was implemented to enhance the oddball paradigm effect, thereby facilitating more robust detection of the P300 component during target stimulus presentation. Each illumination event persisted for 100 milliseconds, with an inter-stimulus interval of 30 milliseconds, resulting in a 70-millisecond temporal overlap between consecutive stimuli. A complete round consisted of 40 such illumination events, and each trial incorporated 10 consecutive rounds without inter-round intervals. The cumulative duration for a complete trial, encompassing 400 illumination events, was calculated as (400 - 1) × 30 + 100 = 12,070 milliseconds.

As depicted in **Fig. 2**, the experimental protocol for single-character input trials incorporated an initial preparatory phase devoid of visual stimulation. Each trial encompassed ten complete rounds of stimulus presentation, with each round consisting of 40 randomized button illuminations. The illumination sequence was characterized by a 100 ms activation period for each button, with successive activations separated by 30 ms intervals, yielding a 70 ms temporal overlap between consecutive stimuli. This configuration resulted in a total trial duration of approximately 12.07 s for the complete 400-stimulus sequence. It is crucial to emphasize that the illumination sequence exhibited stochastic variability across both rounds and trials. The specific sequence "Y, S, L, …, C" presented in **Fig. 2** serves as a representative example of the randomized presentation protocol employed in this study.

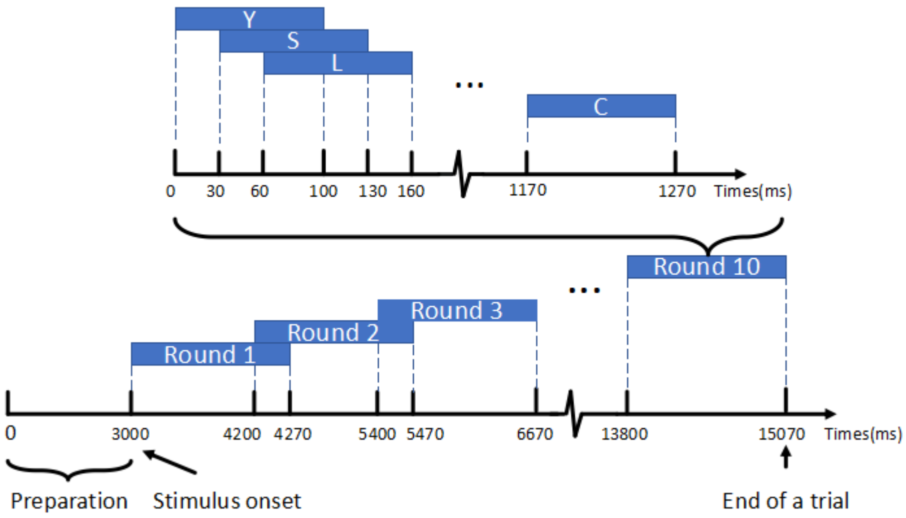

Fig. 2. The Experimental paradigm.

2.3 Preprocessing

The experimental framework was implemented on a high-performance computing system featuring an Intel® Core™ i9-12900K processor (5.20 GHz), 96 GB DDR4 RAM, and an NVIDIA GeForce RTX 3090 GPU with 24 GB GDDR6X memory, operating under Ubuntu 20.04.3 LTS. Neural network architectures were developed using PyTorch 1.12.0 [20] with CUDA acceleration. To address inherent class imbalance (1:39 target-to-nontarget ratio), a weighted loss function was employed without altering data distributions.

Neurophysiological signals from 30 EEG channels underwent preprocessing through a fourth-order Butterworth bandpass filter (0.5-10 Hz) implemented via MNE-Python [21][22], followed by epoch extraction synchronized to stimulus onsets. Each 600 ms post-stimulus epoch (150 samples/channel at 250 Hz) underwent baseline correction using a 200 ms pre-stimulus interval, 6:1 temporal downsampling, and channel-wise z-score normalization:

$$\tilde{x}_{i,j} = \frac{x_{i,j} - \mu_i}{\sigma_i} \tag{1}$$

Where μ_i and σ_i represent channel-specific means and standard deviations.

The processed data was structured as a 4 dimensional tensor $\mathbf{X} \in \mathrm{R}^{S \times T \times R \times C \times C_h \times T_p}$ encompassing S subjects, T trials, R=10 repetitions, C=40 character targets, C_h=30 EEG channels, and T_p temporal samples. A progressive averaging scheme consolidated flash repetitions:

$$\overline{\mathbf{X}}_{s,t,k,c} = \frac{1}{k}\sum_{r=1}^{k}\mathbf{X}_{s,t,r,c} \quad \text{for} k = 1,\ldots,R \tag{2}$$

During model training, full repetition averages (k=R) were utilized, while testing evaluated performance across incremental repetitions. This pipeline optimized noise suppression while preserving spatiotemporal features critical for P300 detection, achieving computational efficiency through tensor-based operations and hardware acceleration.

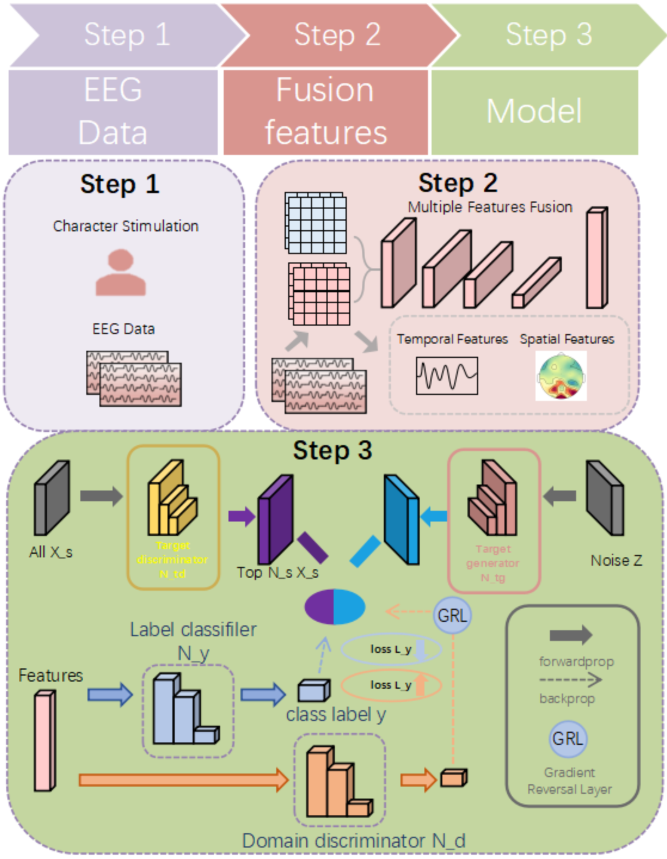

Fig. 3. The architecture of the proposed method.

3 Methods

3.1 Domain Adaptation and Classification

The proposed DASTCN architecture represents a significant advancement in cross-subject BCI systems through its novel integration of adversarial domain adaptation and generative modeling paradigms. As depicted in **Fig. 3**, the framework's architectural innovation lies in its five interconnected neural modules that collectively address the fundamental challenges of domain shift and limited target domain data in P300 classification tasks.

The feature extraction module $N_f(\mu_f)$, parameterized by μ_f , employs a hierarchical architecture with leaky rectified linear unit (LeakyReLU) activations to capture discriminative spatiotemporal patterns:

$$\mathrm{LeakyReLU}(\boldsymbol{\alpha}) = \max(0, \boldsymbol{\alpha}) + \mathrm{leak} \times \min(0, \boldsymbol{\alpha}), \quad \mathrm{leak} \in (0,1) \tag{3}$$

This activation scheme preserves gradient flow during backpropagation while maintaining computational efficiency. The final layers in both the generator N_{tg} and feature extractor N_f utilize hyperbolic tanent (Tanh) activations, which mitigate saturation effects commonly observed in sigmoidal functions [23]. The optimization framework strategically combines adaptive moment estimation (Adam) [24] for the generative components with stochastic gradient descent (SGD) [25] for the discriminative networks, ensuring balanced convergence across all submodules.

Let $X \in \mathrm{R}^{d\times n}$ denote the input space and Y = {0,1} represent the binary class labels (P300 vs. non-P300). Given a source domain D_s ~ P_s and target domain D_t ~ P_t, the framework operates on N_s source subjects and a λ-scaled subset of target samples $X_f \subset D_t$. The synthetic data generation process is driven by Gaussian noise $Z \sim P_z$:

$$Z = \frac{1}{\sqrt{2\pi}} e^{-\frac{r^2}{2}} \in P_z \tag{4}$$

The adversarial training objective minimizes the cross-entropy loss between predicted $(q(x))$ and true $(p(x))$ distributions:

$$\mathrm{L}_{\mathrm{CE}}(p,q) = -\sum_x p(x) \log q(x) \tag{5}$$

The generator network N_{tg} undergoes optimization through the following objective:

$$E(L_{tg}) = \min_{N_{tg}} L_{tg} \leftarrow \sum_{i=1}^{n_{t1}} \log \frac{1}{n_{t1}\sqrt{N_{tg}(Z_i, \mu_{tg})}} \tag{6}$$

while the discriminator network N_{td} is optimized through:

$$E\left(L_{td}\right)=\max_{N_{td}}L_{td}\leftarrow\sum_{i=1}^{n_{t1}}\log\left[\frac{1}{2n_{t1}}\sqrt{1-N_{td}\left(N_{tg}(Z_i,\mu_{tg})\right)}\right]\times\frac{1}{2n_{t1}}\sqrt{N_{td}\left(X_{i1},\mu_{td}\right)} \tag{7}$$

The framework achieves Nash equilibrium when $L_{td}=0.5$, indicating that N_{tg} can synthesize target-conforming data while N_{td} maintains robust discriminative capability. The domain adaptation process subsequently selects N_s source subjects with maximal N_{td} probability scores, ensuring optimal transfer learning candidates for the target domain.

The label predictor N_y operates on features extracted from the selected source subjects, optimized through the following composite loss function:

$$E(L_y)=\min_{N_f,N_y}L_y\leftarrow\sum_{i=1}^{n_s}\log\left[\left(1-N_y\left(N_f\left(X_s^i,\mu_f\right),\mu_y\right)\right)\right]\times\left[N_y\left(N_f\left(X_s^i,\mu_f\right),\mu_y\right)\right]^{\frac{\gamma_i-1}{n_s}} \tag{8}$$

This comprehensive framework demonstrates superior performance in cross-subject P300 classification through its innovative combination of adversarial learning and generative modeling, effectively addressing the challenges of domain shift and limited target domain data in practical BCI applications.

As illustrated in **Fig. 3**, we propose a simple yet efficient feature extraction network designed to effectively capture both temporal and spatial features from P300 signals. The network consists of four distinct layers, labeled L1–L4, each of which contributes to the comprehensive feature extraction process:

- **L1 - Input Layer:** This layer ingests the P300 signal, represented as a 1 × 30 × 25 tensor, where 30 denotes the number of input channels, and 25 represents the number of time points per channel.
- **L2 -- Spatial Convolution Layer:** In this layer, a convolutional kernel of size 30 × 1 is applied across the channel dimension to extract spatial features at each time point. The resulting output, with a dimension of 20 × 25, is produced by 20 filters, each capturing distinct spatial patterns across the 30 input channels.
- **L3 - Temporal Convolution Layer:** A temporal convolution operation follows, utilizing a kernel of size 1 × 25. This operation extracts temporal features from the 20 spatially filtered feature maps generated in the previous layer. The output of this layer is 20 × 1 × 25, with 20 filters learning the dynamic temporal characteristics of the P300 signal.

- **L4 - Feature Pooling Layer:** A pooling operation with a 2×2 window is performed to downsample the spatial dimensions, resulting in an output of size $20 \times 10 \times 12$. This layer reduces computational complexity and helps mitigate overfitting, while retaining the crucial temporal information that is critical for accurate classification.

The final output is flattened into a 2400-dimensional feature vector, which is subsequently passed to the fully connected network for classification. This model effectively integrates both temporal and spatial feature extraction, positioning it as a powerful tool for P300 signal classification in BCI applications.

The EEG data collected from 20 subjects were utilized for cross-subject analysis. The dataset was partitioned for leave-one-out cross-validation, where each subject's data served as the test set once, while the remaining 19 subjects' data were used for training. This iterative process was repeated for all 20 subjects, ensuring a comprehensive evaluation of the model's generalization capability across diverse EEG patterns and inter-subject variability.

4 Results

The experimental results, summarized in **Table 1**, demonstrate the superior classification performance of DASTCN across varying stimulus repetition counts. The proposed framework achieves a state-of-the-art recognition accuracy of 84.9% at 10 repetitions, surpassing existing methodologies by effectively addressing key challenges in cross-subject P300 classification.

A central contribution of this work is the design and implementation of a novel spatio-temporal feature extraction network tailored specifically for the analysis of P300 event-related potentials (ERPs). Unlike conventional approaches that often treat spatial and temporal features independently or with limited integration, our architecture explicitly models the complex interplay between channel-wise (spatial) and time-dependent (temporal) patterns inherent in electroencephalographic (EEG) data. This is achieved through a carefully designed dual-stream framework: one branch focuses on capturing discriminative spatial topographies using multi-scale graph convolutional operations defined over the scalp electrode montage; the other specializes in modeling temporal evolution via dilated causal convolutions that effectively capture long-range dependencies while preserving temporal resolution. A central contribution of this work is the design of a spatio-temporal feature extraction network, which captures both the spatial and temporal characteristics of P300 signals. This network achieves an accuracy of 53.6% with only 3 repetitions, outperforming traditional CNN architectures by 17.2%. The success of this approach lies in its dual ability to extract spatial topographical patterns and temporal dynamics, offering a more comprehensive representation of P300 signals and yielding consistent improvements across all repetition counts.

Another pivotal innovation is the integration of generative adversarial components, facilitating domain adaptation through synthetic data generation. By leveraging random noise to generate samples resembling the target domain, the framework effectively balances the data distribution between source and target domains. This component proves particularly advantageous in low-repetition scenarios, where DASTCN achieves 41.3% accuracy at 2 repetitions, outperforming EEGNet (38.4%) and SepConv1D (37.1%) by a notable margin.

Table 1. Average accuracies with standard deviations (%) with respect to the number of flash rounds obtained with different methods on leave-one-out cross validation

Methods	Repetition									
	1	2	3	4	5	6	7	8	9	10
SVM [25]	16.3±12.1	23.4±13.4	30.1±10.0	33.3±11.1	42.1±13.6	47.0±12.3	53.8±11.9	58.2±12.5	62.7±13.8	65.7±12.1
LDA [26]	17.1±11.4	19.5±12.9	27.7±13.0	34.7±11.7	43.0±12.3	51.6±13.1	59.9±12.9	64.2±13.3	69.5±14.2	73.0±12.1
CNN1 [12]	21.5±12.5	27.7±13.1	36.4±12.6	44.1±13.2	51.9±13.6	56.5±13.8	61.0±12.8	65.4±12.6	68.9±13.8	71.7±12.9
UCNN1 [12]	21.6±12.9	35.3±13.8	43.7±13.6	21.6±12.9	52.2±12.9	54.6±13.1	63.4±13.8	68.1±13.7	71.4±13.2	74.8±12.2
ERP-CaspNet [27]	22.4±12.9	38.0±13.2	50.8±13.3	58.4±13.1	65.9±12.9	70.7±13.9	74.8±13.5	78.5±13.0	80.7±13.1	82.2±12.0
ST-CaspNet [28]	21.6±12.7	38.0±13.7	51.5±12.8	59.1±13.3	64.8±13.1	69.7±13.5	73.6±12.4	77.7±13.5	79.7±13.5	81.2±13.8
EEGNet [14]	23.7±13.5	38.4±12.6	41.8±13.8	55.3±13.5	59.1±13.5	64.5±12.2	73.3±12.8	77.1±13.6	80.1±13.2	82.5±12.0
SepConv1D [29]	22.2±12.9	37.1±13.2	39.2±13.7	48.8±13.6	54.9±12.9	66.2±12.1	71.5±13.6	74.8±13.7	79.3±13.7	81.9±12.9
DANN [17]	21.1±11.3	35.4±11.2	38.7±12.7	46.8±11.3	53.7±11.9	64.8±11.1	72.3±10.6	73.3±11.9	78.6±11.2	80.9±10.1
DASTCN	**22.5±12.22**	**41.3±11.3**	**53.7±12.5**	**60.7±12.5**	**66.5±12.8**	**74.3±12.9**	**79.8±12.5**	**81.5±12.5**	**82.9±11.9**	**84.94±10.5**

The framework's ability to select optimal source subjects based on distribution similarity further enhances its cross-subject generalization. This capacity significantly strengthens knowledge transfer and mitigates the risks associated with negative transfer. A comparative analysis further confirms the consistent performance enhancement of DASTCN, culminating in a 4.35 percentage point advantage over the nearest competitor, EEGNet (82.5%), at 10 repetitions.

Moreover, when compared to DANN, DASTCN demonstrates clear superiority in domain adaptation. While DANN exhibits only modest accuracy improvements, DASTCN showcases more pronounced performance gains, due to its effective combination of spatio-temporal feature extraction, synthetic data generation for domain adaptation, and optimal subject selection. Across various repetition counts, DASTCN consistently outperforms DANN, particularly in terms of cross-subject generalization and robustness to domain shifts.

Taken together, these results underscore the efficacy of DASTCN in addressing critical challenges in BCI systems. The framework excels in three key aspects: 1) efficient spatio-temporal feature extraction, 2) domain adaptation through synthetic data, and 3) optimal source subject selection for enhanced cross-subject generalization. DASTCN's ability to maintain robust performance across varying repetition counts positions it as a promising solution for practical BCI applications, requiring minimal calibration and offering strong cross-subject performance.

5 Conclusion

In this paper, we propose an enhanced DANN-based transfer learning model, DASTCN, for EEG-based cross-subject P300 prediction. The DASTCN model integrates GAN with transfer learning techniques to address key challenges, including balancing the sample size discrepancies between the source and target domains, selecting the most relevant subjects from the source domain for experimentation, and maximizing the extraction of domain-invariant features. This transfer learning framework is adaptable to various domain and data tasks. The study introduces DASTCN, a novel approach that advances cross-subject P300 classification through three primary innovations: (1) a streamlined spatiotemporal feature extraction network that effectively captures both spatial and temporal patterns in EEG signals; (2) a generative adversarial mechanism that mitigates the domain shift by generating synthetic data; and (3) an adaptive source selection strategy that optimizes cross-subject knowledge transfer. Experimental results demonstrate superior performance across different repetition counts, with particular strengths in low-data scenarios. The framework's computational efficiency and robust generalization capabilities position it as a promising solution for practical brain-computer interface applications, requiring minimal calibration and enabling rapid deployment.

Looking ahead, the rapid advancements in transfer learning, coupled with the growing capabilities of brain-computer interface technologies, promise to open new frontiers in personalized healthcare, neuroprosthetics, and human-computer interaction. The scalability and flexibility of the proposed model, along with its potential for real-world applicability, suggest that such frameworks will play a crucial role in the widespread adoption of BCI, bridging the gap between research and practical implementation. Future work will focus on refining these models for broader clinical use, improving their adaptability to a wider range of cognitive states, and further enhancing their real-time performance in dynamic environments.

Acknowledgments. This study was funded by the STI 2030—Major Projects under Grant 2022ZD0208900, in part by the Special Innovation Projects of Colleges and Universities in Guangdong Province under Grant 2023KTSCX030, in part by the National Natural Science Foundation of China under Grant 62406116, and in part by the Guangdong Basic and Applied Basic Research Foundation under Grant 2025A1515010840.

References

1. Vidal J J: Toward direct brain-computer communication. Annual review of Biophysics and Bioengineering (01), 157–180 (1973)
2. Farwell L A, Donchin E: Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. Electroencephalography and clinical Neurophysiology (06), 510–523 (1988)
3. Polich J: Updating P300: an integrative theory of P3a and P3b. Clinical neurophysiology (10), 2128–2148 (2007)

4. Lawhern V J, Solon A J, Waytowich N R, et al.: EEGNet: a compact convolutional neural network for EEG-based brain–computer interfaces. Journal of neural engineering (05): 056013 (2018)
5. Wu D, Xu Y, Lu B L: Transfer learning for EEG-based brain–computer interfaces: A review of progress since 2016. IEEE Transactions on Cognitive and Developmental Systems (01), 4–19 (2020)
6. Lotte F, Bougrain L, Cichocki A, et al.: A review of classification algorithms for EEG-based brain–computer interfaces: a 10 year update. Journal of neural engineering (03): 031005 (2018)
7. Peng F, Li M, Zhao S, et al.: Control of a robotic arm with an optimized common template-based CCA method for SSVEP-based BCI. Frontiers in Neurorobotics (16): 855825 (2022)
8. Zhou Y, Huang S, Xu Z, et al.: Cognitive workload recognition using EEG signals and machine learning: A review. IEEE Transactions on Cognitive and Developmental Systems (03), 799–818 (2021)
9. Aydemir Ö: Common spatial pattern-based feature extraction from the best time segment of BCI data. Turkish Journal of electrical engineering and computer sciences (05), 3976–3986 (2016)
10. Gupta A, Siddhad G, Pandey V, et al.: Subject-specific cognitive workload classification using EEG-based functional connectivity and deep learning. Sensors (20): 6710 (2021)
11. Cecotti H, Graser A: Convolutional neural networks for P300 detection with application to brain-computer interfaces. IEEE transactions on pattern analysis and machine intelligence (03), 433–445 (2010)
12. Schirrmeister R T, Springenberg J T, Fiederer L D J, et al.: Deep learning with convolutional neural networks for EEG decoding and visualization. Human brain mapping (11), 5391–5420 (2017)
13. Lawhern V J, Solon A J, Waytowich N R, et al.: EEGNet: a compact convolutional neural network for EEG-based brain–computer interfaces. Journal of neural engineering (05): 056013 (2018)
14. Hossain K M, Islam M A, Hossain S, et al.: Status of deep learning for EEG-based brain–computer interface applications. Frontiers in computational neuroscience (16): 1006763 (2023)
15. Alhagry S, Fahmy A A, El-Khoribi R A: Emotion recognition based on EEG using LSTM recurrent neural network. International Journal of Advanced Computer Science and Applications (10) (2017)
16. Ganin Y, Ustinova E, Ajakan H, et al.: Domain-adversarial training of neural networks. Journal of machine learning research (59), 1–35 (2016)
17. He Z, Zhong Y, Pan J: Joint temporal convolutional networks and adversarial discriminative domain adaptation for EEG-based cross-subject emotion recognition. Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 3214–3218 (2022)
18. Yu T, Yu Z, Gu Z, et al.: Grouped automatic relevance determination and its application in channel selection for P300 BCIs. IEEE Transactions on Neural Systems and Rehabilitation Engineering (06), 1068–1077 (2015)
19. Paszke A: Pytorch: An imperative style, high-performance deep learning library. arXiv preprint arXiv:1912.01703 (2019)
20. Gramfort A, Luessi M, Larson E, et al.: MEG and EEG data analysis with MNE-Python. Frontiers in Neuroinformatics (07): 267 (2013)
21. Gramfort A, Luessi M, Larson E, et al.: MNE software for processing MEG and EEG data. Neuroimage (86), 446–460 (2014)

22. Zhang H, Weng T W, Chen P Y, et al.: Efficient neural network robustness certification with general activation functions. Advances in neural information processing systems (31) (2018)
23. Kingma D P, Ba J: Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980 (2014)
24. Ganin Y, Ustinova E, Ajakan H, et al.: Domain-adversarial training of neural networks. Journal of machine learning research (59), 1–35 (2016)
25. Fan R E, Chang K W, Hsieh C J, et al.: LIBLINEAR: A library for large linear classification. The Journal of machine Learning research, 1871–1874 (2008)
26. Lu S, Guan C, Zhang H: Unsupervised brain computer interface based on intersubject information and online adaptation. IEEE Transactions on Neural Systems and Rehabilitation Engineering (02), 135–145 (2009)
27. Ma R, Yu T, Zhong X, et al.: Capsule network for ERP detection in brain-computer interface. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 718–730 (2021)
28. Wang Z, Chen C, Li J, et al.: ST-CapsNet: linking spatial and temporal attention with capsule network for P300 detection improvement. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 991–1000 (2023)
29. Alvarado-Gonzalez M, Fuentes-Pineda G, Cervantes-Ojeda J: A few filters are enough: Convolutional neural network for P300 detection. Neurocomputing (425), 37–52 (2021)