Title

MLDA-STG: A Multi-Level Domain Adaptation Spatio-Temporal Graph Network for Cross-Subject EEG Fatigue Detection

Permalink

https://escholarship.org/uc/item/5kg0z89r

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 48(0)

Authors

Zhang, Jiahui

Chen, Rongtao

Huang, Zhuoyi

et al.

Publication Date

2026

MLDA-STG: A Multi-Level Domain Adaptation Spatio-Temporal Graph Network for Cross-Subject EEG Fatigue Detection

Jiahui Zhang[1†], Rongtao Chen[1†], Zhuoyi Huang[1], Chuwen Xie[1], Jiahui Pan (panjiahui@m.scnu.edu.cn)[1] & Lina Qiu (lina.qiu@scnu.edu.cn)[1]

[1]School of Artificial Intelligence, South China Normal University, Foshan, China

Abstract

Reliable decoding of cognitive fatigue from Electroencephalography (EEG) is essential for monitoring sustained attention but is hampered by significant inter-subject variability in neural dynamics. To address the resulting domain shift, we propose the Multi-Level Domain Adaptation Network with an Enhanced Spatio-Temporal Graph (MLDA-STG). Our approach uniquely models the evolving functional connectivity of the brain during cognitive decline through a dynamic graph fusion network, while a collaborative learning mechanism disentangles personalized neural traits from shared fatigue patterns. Crucially, we introduce a Multi-Level Alignment framework that goes beyond global adaptation to align conditional distributions, preserving the semantic structure of cognitive states across individuals. Evaluations on the SEED-VIG and SADT datasets demonstrate that MLDA-STG achieves state-of-the-art performance, offering a robust solution for calibration-free cognitive state assessment. The source code is available at: https://github.com/zjh-sys/MLDA-STG.git.

Keywords: Fatigue detection, electroencephalogram (EEG), cognitive science, domain adaptation, functional connectivity.

Introduction

As a critical component of cognitive state monitoring, driver drowsiness detection serves as a vital instrument for preventing traffic accidents and safeguarding lives. Statistically, fatigue driving is estimated to contribute to as many as 20% of all vehicle collisions (Cui, Yuan, et al., 2023). However, drivers' fatigue is not merely a physiological state of drowsiness, but a profound cognitive failure characterized by the depletion of attentional resources and the degradation of executive functions (Wickens, 2008). From the perspective of neuroergonomics (Parasuraman, 2003), this state represents a critical lapse in sustained attention (vigilance), leading to delayed reaction times and impaired decision-making capabilities. While traditional monitoring relies on overt behavioral cues (e.g., facial expressions), these markers are often downstream effects that manifest only after cognitive decline has occurred and are susceptible to subjective masking. In contrast, Electroencephalography (EEG) has emerged as a gold standard in neuroergonomics and affective computing (F. Wang et al., 2020). By directly measuring Central Nervous System activity, EEG offers a window into the intrinsic neural dynamics underlying vigilance decrements, providing a more objective assessment that captures oscillatory changes preceding behavioral lapses.

However, developing robust EEG-based fatigue detection systems faces two primary obstacles rooted in the complexity of neural information processing:

1) Complexity of Spatio-Temporal Neural Dynamics: The transition from an alert to a drowsy state is not a static shift but a dynamic evolution of neural oscillations coupled across brain regions and time (Zhao et al., 2017). Early research focused on handcrafted features like Power Spectral Density and Differential Entropy fed into Support Vector Machines (Cui, Yuan, et al., 2023; F. Wang et al., 2020). Recently, Deep Learning models, such as Convolutional Neural Networks and Graph Convolutional Networks, have shown capability in decoding these states (Cui, Lan, et al., 2023). Yet, modeling the intricate, non-linear functional connectivity between brain regions—which is essential for capturing the "topology" of fatigue—remains a challenge for standard architectures.

2) Individual Variability and Domain Shift: A fundamental challenge in cognitive modeling is the significant physiological variability between individuals, known as the cross-subject problem. Due to differences in cortical folding, skull conductivity, and baseline neural activity, the EEG distribution for the same cognitive state (e.g., fatigue) varies drastically across subjects (Huang et al., 2025). A model trained on one group (source domain) often fails to generalize to a new, unseen subject (target domain) due to this non-stationarity, formally described as "domain shift." This severely hinders the development of calibration-free systems that can adapt to the heterogeneous neural signatures of the general population.

To mitigate these discrepancies, Domain Adaptation techniques have been introduced to align feature distributions (Huang et al., 2025). Current strategies typically employ distance-based metrics (e.g., Maximum Mean Discrepancy) or adversarial training (Ganin & Lempitsky, 2015; Gretton et al., 2012) to extract domain-invariant features. However, a critical limitation persists: most existing approaches focus predominantly on aligning the global *marginal distribution* of neural features, effectively pulling all data points together. They frequently overlook the alignment of the *conditional distribution*, which encodes the specific semantic relationship between neural patterns and cognitive labels (Feng et al., 2025). This oversight often leads to "negative transfer," where features from different cognitive states (e.g., alert vs. drowsy) are incorrectly clustered together, degrading fine-grained recognition performance.

To address these issues, we propose a novel framework

In G. Boleda, I. Dautriche, M. Maldonado, C. Saldana and T. Torrent (Eds.), *Proceedings of the 48th Annual Conference of the Cognitive Science Society*

named Multi-Level Domain Adaptation Network with Enhanced Spatio-Temporal Graph (MLDA-STG). Our approach is designed to model the brain's dynamic functional connectivity while explicitly bridging the distributional gap between individuals. The main contributions are as follows:

- **Enhanced Spatio-Temporal Graph Information Fusion Network (ESTGF)**: We propose a novel parallel feature extraction network capable of dynamically integrating local, spatial, and temporal neural cues from EEG signals. By modeling the evolution of brain network topology during cognitive decline, this module learns robust, fatigue-related domain-invariant representations.
- **Shared and Personalized Feature Learning**: We develop a collaborative learning mechanism that explicitly disentangles universal neural mechanisms of fatigue from individual physiological variability. This enables the simultaneous capture of cross-subject common cognitive patterns and subject-specific characteristics, balancing generalization with personalization.
- **Multi-Level Alignment Framework (MLA)**: We introduce a multi-stage strategy designed to mitigate inter-subject neural discrepancies. By aligning both marginal and conditional probability distributions of neural features, this framework comprehensively addresses the domain shift problem, ensuring consistent interpretation of fatigue states across different individuals.

Extensive experiments on the SEED-VIG and SADT datasets demonstrate that our model not only achieves state-of-the-art detection performance but also reveals distinct, interpretably compact clusters of cognitive states, offering new insights into the generalized neural markers of human fatigue.

Methods

The overall architecture of the proposed method (MLDA-STG), as illustrated in Fig. 1, consists of three primary components: an enhanced spatio-temporal graph information fusion network (ESTGF) for shared feature extraction, a subject-specific convolutional module for specific feature extraction, and a multi-level alignment (MLA) framework for comprehensive domain adaptation.

Enhanced Spatio-Temporal Graph Information Fusion Network

ESTGF integrates a convolutional neural network backbone, a parallel transformer backbone, and a graph attention network backbone. It dynamically learns and fuses features from three complementary dimensions: local patterns, temporal dependencies, and spatial correlations.

Convolutional Neural Network Backbone The network learns multi-scale local features through initial downsampling and multiple core modules composed of deep convolution and point convolution. The feature maps are converted into feature vectors through global average pooling to achieve multimodal fusion, and reshaped into initial node features for graph attention network processing.

Transformer Backbone A standard transformer encoder is used to capture long-term temporal dependencies (Vaswani et al., 2017). It outputs a global context vector, f_{temp}, which is used for both final feature fusion and to construct dynamic graphs.

Graph Attention Network Backbone To model functional connectivity between different brain regions, each EEG channel is treated as a node in the graph (Veličković et al., 2018), with deep features generated by CNN backbone serving as initial node inputs. Instead of using a fixed adjacency matrix, we design a dynamic graph construction mechanism that generates an adaptive graph structure for each input sample.

First, a learnable base adjacency matrix, $A_{base} \in \mathbb{R}^{N\times N}$ (where N is the number of EEG channels), is defined. A gating matrix, G_m is then generated by passing the temporal context vector f_{temp} through a gating MLP. The dynamic adjacency matrix A_{dyn} is computed as follows:

$$G_m = \mathrm{Sigmoid}(\mathrm{MLP}(f_{temp})), \tag{1}$$

$$A_{dyn} = A_{base} \odot G_m, \tag{2}$$

where $\odot$ denotes element-wise multiplication. This mechanism dynamically adjusts graph connectivity according to the temporal dynamics of the input.

For numerical stability, the dynamic adjacency matrix is symmetrically normalized before graph convolution:

$$\hat{A}_{dyn} = D^{-\frac{1}{2}} A_{dyn} D^{-\frac{1}{2}}, \tag{3}$$

where D is the degree matrix of A_{dyn}.

A multi-head graph attention network then uses self-attention to weigh each node's neighbors and aggregate their spatial information. The feature update for a node i can be summarized as follows:

$$h'_i = \sigma\left(\sum_{j\in N_i} \alpha_{ij} W h_j\right), \tag{4}$$

where h_j is a neighbor node's feature, W is a learnable weight matrix, σ is a non-linear activation, and α_{ij} is the attention coefficient between nodes i and j.

Cross-Attention Fusion Mechanism We fuse three parallel branches using a cross-attention strategy where each branch acts as a query to extract key-value information from the other two. The resulting augmented feature vectors are then concatenated to form the final domain-invariant shared features.

Subject-Specific Convolutional Module

This module is a lightweight convolutional neural network to extract personalized EEG patterns from each subject, complementing shared features. It processes single-frame EEG data to generate a subject-specific feature vector, which is then fused with shared features in an alignment space. The final predictions integrate both shared and subject-specific information, enhancing performance.

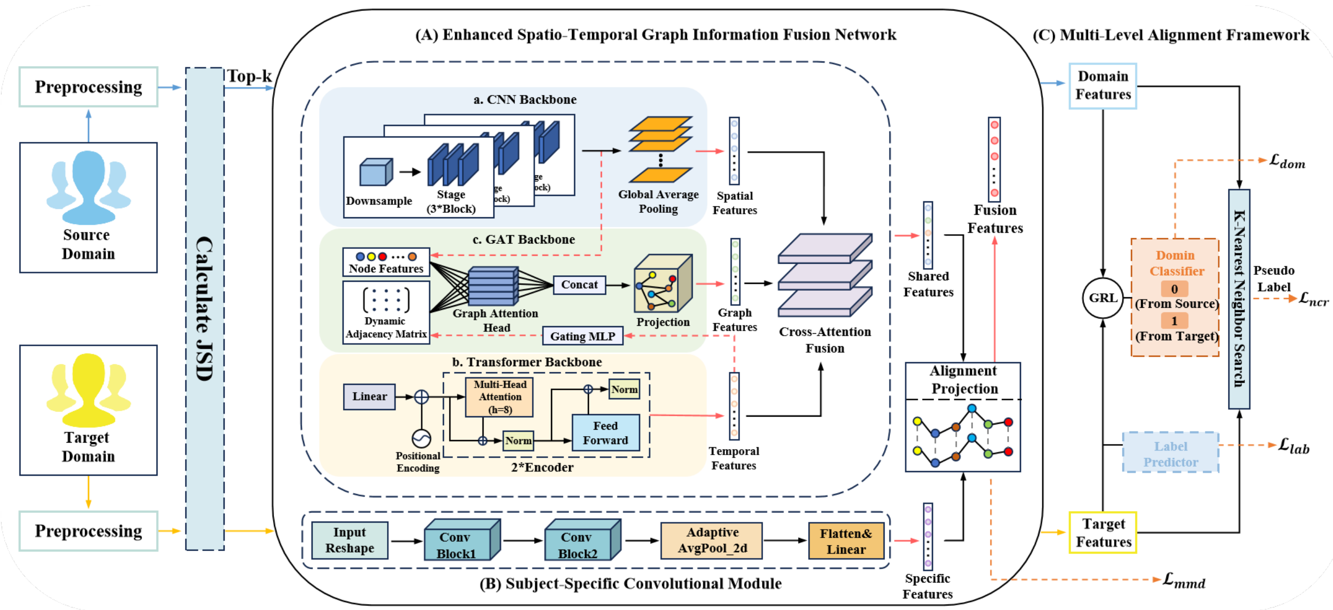

Figure 1: Framework of the proposed multi-level domain adaptation network with enhanced spatio-temporal graph. This framework comprises three key components: (A) an Enhanced Spatio-Temporal Graph Information Fusion Network, (B) a Subject-Specific Convolutional Module, and (C) a Multi-Level Alignment Framework.

Multi-Level Alignment Framework

Multi-Source Selection To reduce transfer noise, we select the top k source domains most similar to the target domain. Similarity is measured by the Jensen-Shannon divergence (JSD) between their feature distributions. The JSD is calculated as follows:

$$JSD(P||Q) = \frac{1}{2}D_{KL}(P||M) + \frac{1}{2}D_{KL}(Q||M), \tag{5}$$

where P and Q are the probability distributions of the features for two subjects, $M = \frac{1}{2}(P+Q)$, and D_{KL} is the Kullback-Leibler divergence (Kullback & Leibler, 1951).

Multi-Level Adaptive Computing Since aligning only the marginal probability distributions is insufficient to resolve the conditional probability distribution shifts associated with fatigue categories, we employ a multi-level alignment strategy that calibrates the feature space from coarse to fine through three collaborative stages.

Local Regularization Alignment: At this stage, the Maximum Mean Discrepancy loss function (L_{mmd}) is employed to align the "shared features" extracted by the main network with the "personalized features" extracted by subject-specific modules, ensuring that personalized modules learn effective patterns relevant to fatigue tasks.

$$L_{mmd} = \left\| \frac{1}{|F_{sh}|} \sum_{f_{sh} \in F_{sh}} \phi(f_{sh}) - \frac{1}{|F_p|} \sum_{f_p \in F_p} \phi(f_p) \right\|_{\mathcal{H}}^2, \tag{6}$$

where F_{sh} and F_p are the shared and personalized features, respectively. $\phi(\cdot)$ is a kernel function mapping to a Reproducing Kernel Hilbert Space (RKHS) $\mathcal{H}$.

Global Distribution Alignment: To eliminate domain shifts, this phase employs domain adversarial training (Ganin & Lempitsky, 2015). It minimizes a domain classification loss using a gradient reversal layer, aligning the edge feature distributions across subjects.

$$L_{dom} = -\frac{1}{|D_s| + |D_t|} \sum_{x_i \in D_s \cup D_t} \left[d_i \log(\hat{d}_i) + (1-d_i)\log(1-\hat{d}_i) \right], \tag{7}$$

where D_s and D_t are the source and target domains, respectively. d_i is the true domain label of sample x_i (e.g., 0 for source, 1 for target), and $\hat{d}_i$ is the domain classifier's predicted probability of the domain label.

Neighborhood Consistency Regularization: This stage employs pseudo-labeling to align the conditional probability distribution of the target domain, thereby enhancing classification performance. The soft pseudo-label $\tilde{y}_j$ is calculated as:

$$\tilde{y}_j = \sum_{i \in \mathcal{N}_j} w_{ji} y_i^s, \tag{8}$$

where $\mathcal{N}_j$ is the set of K-nearest source domain neighbors for target sample j, w_{ji} is the distance-based weight for neighbor i, and y_i^s is the true label of neighbor i. The final loss is:

$$L_{ncr} = \frac{1}{|D_t|} \sum_{j \in D_t} (\hat{y}_j^t - \tilde{y}_j)^2, \tag{9}$$

where $|D_t|$ is the number of target domain samples, $\hat{y}_j^t$ is the model's prediction for target sample j, and $\tilde{y}_j$ is its corresponding soft pseudo-label.

Therefore, the overall loss function is defined as:

$$L = \alpha L_{lab} + \beta L_{dom} + \lambda L_{mmd} + \gamma L_{ncr}, \qquad (10)$$

where L_{lab} is the supervised classification loss on the source domain, and α, β, λ, and γ are weights that balance the different alignment stages.

Experiments and Results

EEG Datasets and Pre-Processing

In this study, we evaluate the proposed framework on two publicly available benchmark datasets for driver drowsiness detection: the SEED-VIG dataset (Zheng & Lu, 2017) and the SADT dataset (Cao et al., 2019). These datasets represent different driving scenarios and data acquisition protocols, providing a comprehensive testbed for cross-subject fatigue detection.

SEED-VIG Dataset The SEED-VIG dataset, collected by the BCMI laboratory, involves 23 sessions from 12 subjects participating in a simulated monotonous driving task. The experiment was conducted in a virtual reality simulator to induce fatigue naturally over a duration of approximately two hours per session. EEG signals were recorded using a Neuroscan system with 18 electrodes positioned according to the standard 10-20 system. Simultaneously, the subjects' eye movements were monitored using SMI eye-tracking glasses. The vigilance level was annotated based on the Percentage of Eye Closure (PERCLOS) index. Following the protocol in, samples were labeled as "Alert" when the PERCLOS value was lower than 0.35 and as "Drowsy" when it exceeded 0.7. Data segments falling between these thresholds were discarded to ensure clear class separability.

SADT Dataset The SADT (Sustained-Attention Driving Task) dataset focuses on event-related fatigue responses. It comprises data from 11 subjects who performed a driving task in a virtual simulator where random lane-departure events were introduced. Subjects were required to correct the vehicle's trajectory immediately after a deviation occurred. EEG signals were captured using a Scan SynAmps2 Express system with 30 EEG electrodes and 2 reference electrodes. The fatigue state was determined by the subject's reaction time to the lane-departure events. Longer reaction times indicated a higher level of drowsiness. For our experiments, we extracted 3-second EEG epochs prior to the deviation events to capture the mental state preceding the response.

Data Pre-Processing To ensure consistency across datasets and remove artifacts, a rigorous pre-processing pipeline was applied. The raw EEG signals from both datasets were downsampled to 128 Hz to reduce computational complexity. A band-pass filter (1–50 Hz) was applied to remove high-frequency noise and DC drift. We utilized Differential Entropy (DE) as the core feature representation (Duan et al., 2013), as it has been proven effective in distinguishing EEG patterns related to fatigue. The continuous EEG signals were decomposed into five standard frequency bands: Delta (1–4 Hz), Theta (4–8 Hz), Alpha (8–14 Hz), Beta (14–31 Hz), and Gamma (31–50 Hz). The DE features were calculated for each frequency band, resulting in a comprehensive feature vector that characterizes the energy distribution of the brain signals across different rhythms. After data preprocessing (including filtering, downsampling, and sample segmentation) and extraction of differential entropy as the core feature, both datasets are utilized for a binary classification task.

Cross-Subject Experiment

We employ a rigorous Leave-One-Subject-Out (LOSO) (Lotte et al., 2007) cross-validation strategy for evaluation. Specifically, data from one subject is selected as the target domain each time, while data from all other subjects constitutes the source domain. This process is repeated for each subject, with evaluation ultimately based on average performance metrics.

Experimental Settings We introduced the multi-source domain selection algorithm proposed in Section Multu-Source Selection of the Methods section to reduce migration noise caused by individual variations. Before training commenced, we computed the JSD values between the target subject's feature distribution and that of each subject in the source domain candidate pool. We select the $K = 10$ subjects with the smallest JSD values as the source domain for that training round. We trained the model for 1000 epochs on an NVIDIA RTX 4090 GPU using the Adam optimizer (weight decay: 1×10^{-4}). The initial learning rates were set to 1×10^{-4} for the feature extractor/classifier and 1×10^{-5} for the discriminator, decaying to 1×10^{-6} via a cosine annealing schedule.

Baselines To comprehensively evaluate the performance of our proposed MLDA-STG, we compared it against six state-of-the-art methods, categorized into advanced deep learning architectures and domain adaptation frameworks:

- **EEG-Conformer** (Song et al., 2023): A hybrid architecture that integrates convolutional neural networks with Transformer encoders to capture both local temporal features and global dependencies in EEG signals for decoding and visualization.
- **InterpretCNN** (Cui, Lan, et al., 2023): An interpretable convolutional neural network designed for cross-subject drowsiness recognition, employing multi-scale convolutions to extract discriminative features from single-channel EEG.
- **TFORMER** (Li et al., 2024): A time-frequency Transformer framework that incorporates batch normalization strategies to model the complex multi-dimensional dependencies of EEG signals for fatigue recognition.
- **ICNN+FGloWD** (Li et al., 2023): An enhanced ensemble learning framework utilizing deep random vector functional link networks, designed to improve the robustness and stability of driver fatigue recognition.

- **DAN** (S. Wang et al., 2024): An improved deep residual network approach initially designed for complex fault diagnosis, which we adopted as a robust deep learning baseline to evaluate feature extraction capability under imbalanced data conditions.
- **SMLDA** (Huang et al., 2025): A multi-level domain adaptation framework specifically designed for EEG fatigue detection, which improves generalization by aligning marginal and conditional distributions across subjects.

Experimental Results Some algorithms are used to compare and the experimental results are depicted in Table 1.

Table 1: Performance Comparison of MLDA-STG and State-of-the-Art Methods on SEED-VIG and SADT Datasets Using Leave-One-Subject-Out (LOSO) Cross-Validation.

Datasets	Methods	P_{acc} (%)	F1 score (%)
SEED-VIG	EEG-Conformer	81.70 ± 06.44	81.94 ± 05.77
	InterpretCNN	90.31 ± 05.77	90.14 ± 05.63
	ICNN+FGloWD	91.80 ± 04.03	91.75 ± 03.87
	TFORMER	91.20 ± 04.46	91.28 ± 06.25
	DAN	92.29 ± 04.69	92.10 ± 03.91
	SMLDA	94.20 ± 02.27	94.62 ± 03.24
	Ours	**95.37 ± 03.11**	**95.41 ± 06.32**
SADT	EEG-Conformer	73.10 ± 08.01	73.37 ± 07.92
	InterpretCNN	77.85 ± 07.91	77.75 ± 06.88
	ICNN+FGloWD	79.40 ± 09.46	79.43 ± 09.22
	TFORMER	79.70 ± 06.16	79.60 ± 07.19
	DAN	80.99 ± 08.13	80.72 ± 07.47
	SMLDA	84.30 ± 07.48	84.39 ± 08.21
	Ours	**91.39 ± 07.13**	**91.34 ± 04.53**

Note: P_{acc} represents Classification Accuracy; F1 score is the harmonic mean of Precision and Recall.

Table 1 shows the comparison of fatigue recognition accuracy between our model and other models on both datasets. The powerful feature extractor and comprehensive alignment strategy work in synergy, enabling it to capture rich spatiotemporal dynamic information and effectively correct shifts in marginal and conditional distributions.

Ablation Experiment

To comprehensively validate the contribution of each key component in our proposed architecture, we performed an ablation study by systematically removing individual modules. We designed five distinct ablation variants to verify specific design choices: (1) **w/o SSCM**: This variant removes the Subject-Specific Convolutional Module to validate the necessity of capturing personalized feature patterns. (2) **w/o GAT**: This variant discards the Graph Attention Network from the shared feature extractor, assessing the model's ability to capture dynamic spatial dependencies without graph topology. (3) **w/o LRA**: This variant excludes the Local Regularization Alignment, testing the importance of aligning shared and personalized feature branches. (4) **w/o MSS**: This variant disables the Multi-Source Selection strategy (based on JSD), utilizing all available source subjects regardless of distribution similarity to test the impact of transfer noise. (5) **w/o NCR**: This variant removes the Neighborhood Consistency Regularization, evaluating the role of conditional distribution alignment in fine-grained classification. The experimental results are reported in Table 2.

Table 2: Results of the ablation experiment.

Datasets	Methods	P_{acc} (%)	F1 score (%)
SEED-VIG	w/o SSCM	87.43 ± 06.18	87.52 ± 07.79
	w/o GAT	83.44 ± 05.09	83.53 ± 05.01
	w/o LRA	88.10 ± 07.54	88.11 ± 06.42
	w/o MSS	91.71 ± 05.71	91.88 ± 07.70
	w/o NCR	85.50 ± 06.71	85.37 ± 06.36
	Ours	**95.37 ± 03.11**	**95.41 ± 06.32**
SADT	w/o SSCM	79.35 ± 09.01	79.43 ± 08.81
	w/o GAT	78.45 ± 08.85	78.28 ± 08.50
	w/o LRA	84.37 ± 07.77	84.25 ± 06.24
	w/o MSS	76.59 ± 07.69	76.65 ± 06.93
	w/o NCR	81.80 ± 06.70	81.82 ± 08.35
	Ours	**91.39 ± 07.13**	**91.34 ± 04.53**

Note: P_{acc} represents Classification Accuracy; F1 score is the harmonic mean of Precision and Recall.

Through ablation experiments, the critical roles of each module within MLDA-STG were validated. Table 2 demonstrates that at the feature extraction level, the combination of the personalized feature extractor and the enhanced spatiotemporal graph attention network extracts robust domain-invariant features. At the domain adaptation level, the MLA strategy effectively addresses the conditional probability distribution bias in fatigue classification, ensuring efficient knowledge transfer.

Discussion and Visualization

To provide a comprehensive understanding of how the proposed MLDA-STG framework achieves superior performance in cross-subject driver drowsiness detection, we conduct a multi-level visualization analysis. This analysis covers latent feature distributions, topological structural alignment, and neurophysiological interpretability.

Feature Distribution Alignment As illustrated in the t-SNE visualization (**Fig. 2**), the latent features extracted by our model exhibit high intra-class compactness and clear inter-class separation. Notably, after the multi-level domain adaptation process, features from the source and target domains are effectively aligned in the shared manifold. This indicates that MLDA-STG successfully filters out subject-specific interference and captures domain-invariant representations, which is the primary reason for the observed accuracy improvements on the SEED-VIG and SADT datasets.

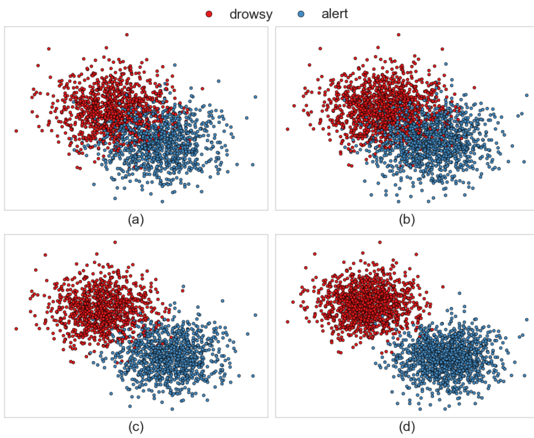

Figure 2: t-SNE visualization of the MLDA-STG model's feature distributions. The subfigures show features on the SEED-VIG (a, c) and SADT (b, d) datasets, comparing them before (a, b) and after (c, d) training.

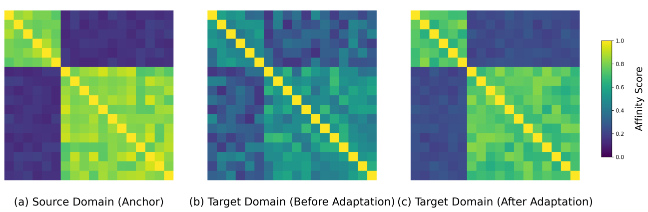

Figure 3: Comparison of functional connectivity matrices on SEED-VIG dataset: (a) Source Domain (Anchor), (b) Target subject before adaptation, and (c) Target subject after adaptation. Note: Electrode labels are omitted for visual clarity; the 18 electrodes are ordered to reflect intrinsic functional clusters (e.g., temporal and posterior areas).

Topological Structural Consistency To investigate the model's capability in aligning spatial structures across different subjects, we visualize the functional connectivity matrices in three stages (**Fig. 3**). Initially, due to significant individual variations, the target subject exhibits fragmented and inconsistent connectivity patterns before adaptation (Fig. 3b). However, after applying the proposed MLDA-STG framework, the target topology effectively bridges the domain discrepancy and converges toward a structured, block-diagonal pattern (Fig. 3c), which closely resembles the intrinsic functional dependencies of the source anchor (Fig. 3a). This evolution confirms that our model successfully reconstructs global spatial dependencies and extracts robust, domain-invariant representations from heterogeneous individual signals. Such structural alignment ensures the stability of the classification manifold, providing a solid foundation for high-performance cross-subject fatigue detection.

Neurophysiological Interpretability To further validate the spatial interpretability of the MLDA-STG, we visualize the brain topography maps for both Alert and Drowsy states

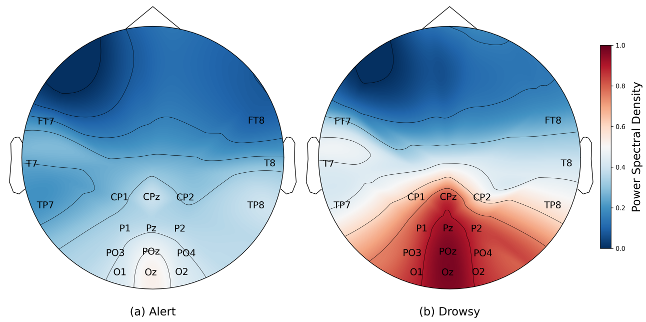

Figure 4: Alpha-band topographic maps for Alert (a) and Drowsy (b) states on the SEED-VIG dataset. Data labels were determined based on the PERCLOS index threshold (< 0.35 for Alert and > 0.7 for Drowsy).

(**Fig. 4**). As illustrated, a significant increase in Alpha-band power is observed in the parietal and occipital regions (e.g., P_z, O_z) during the drowsy state, which aligns with established neurophysiological findings (Zheng & Lu, 2017). These results demonstrate that our model effectively captures discriminative spatial patterns that have clear physiological significance, rather than merely relying on stochastic correlations. This groundedness in neurobiology further justifies the robustness and reliability of our framework in real-world driving scenarios.

Conclusion

In this paper, we proposed a multi-level domain adaptation network with an enhanced spatio-temporal graph to address the challenge of cross-subject domain shift in EEG-based cognitive fatigue detection. Beyond simple feature extraction, our model dynamically extracts and fuses discriminative spatio-temporal features to model the brain's intrinsic functional connectivity, effectively capturing the complex neural reorganizations inherent to vigilance decrements. Collaboratively, it learns to capture personalized neural patterns while aligning feature distributions from coarse to fine levels, ensuring the preservation of the semantic structure of cognitive states across subjects. Comprehensive experiments on two public datasets, including detailed ablation studies, validate our model's effectiveness and superior performance against existing methods. These results confirm that our work provides a robust solution for developing practical, calibration-free EEG fatigue detection systems, paving the way for reliable real-time cognitive state monitoring in neuroergonomics.

Acknowledgments

This work was supported by the National Natural Science Foundation of China under grants 62576142 and 82302339, the Guangdong Basic and Applied Basic Research Foundation under grant 2024A1515010524, and the Characteristic Innovation Project of Colleges and Universities in Guangdong Province under grant 2024KTSCX097.

References

Cao, Z., Chuang, C.-H., King, J.-K., & Lin, C.-T. (2019). Multi-channel EEG recordings during a sustained-attention driving task. *Scientific Data*, *6*(1), 19.

Cui, J., Lan, Z., Sourina, O., & Müller-Wittig, W. (2023). EEG-based cross-subject driver drowsiness recognition with an interpretable convolutional neural network. *IEEE Transactions on Neural Networks and Learning Systems*, *34*(10), 7921–7933.

Cui, J., Yuan, L., Li, R., Wang, Z., Yang, D., & Jiang, T. (2023). Benchmarking EEG-based cross-dataset driver drowsiness recognition with deep transfer learning. *2023 45th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, 1–6.

Duan, R.-N., Zhu, J.-Y., & Lu, B.-L. (2013). Differential entropy feature for EEG-based emotion classification. *2013 6th International IEEE/EMBS Conference on Neural Engineering (NER)*, 81–84.

Feng, X., Dai, S., & Guo, Z. (2025). Pseudo-label-assisted subdomain adaptation network with coordinate attention for EEG-based driver drowsiness detection. *Biomedical Signal Processing and Control*, *101*, 107132.

Ganin, Y., & Lempitsky, V. (2015). Unsupervised domain adaptation by backpropagation. *Proceedings of the 32nd International Conference on Machine Learning*, *37*, 1180–1189.

Gretton, A., Borgwardt, K. M., Rasch, M. J., Schölkopf, B., & Smola, A. (2012). A kernel two-sample test. *Journal of Machine Learning Research*, *13*(25), 723–773.

Huang, F., Wang, Q., Chen, L., Mei, W., Zhang, Z., & Chen, Z. (2025). Multi-level domain adaptation for improved generalization in electroencephalogram-based driver fatigue detection. *Engineering Applications of Artificial Intelligence*, *142*, 109825.

Kullback, S., & Leibler, R. A. (1951). On information and sufficiency. *The Annals of Mathematical Statistics*, *22*(1), 79–86.

Li, R., Gao, R., Yuan, L., Suganthan, P. N., Wang, L., & Sourina, O. (2023). An enhanced ensemble deep random vector functional link network for driver fatigue recognition. *Engineering Applications of Artificial Intelligence*, *123*, 106237.

Li, R., Hu, M., Gao, R., Wang, L., Suganthan, P., & Sourina, O. (2024). Tformer: A time-frequency transformer with batch normalization for driver fatigue recognition. *Advanced Engineering Informatics*, *62*, 102575.

Lotte, F., Congedo, M., Lécuyer, A., Lamarche, F., & Arnaldi, B. (2007). A review of classification algorithms for EEG-based brain-computer interfaces. *Journal of Neural Engineering*, *4*, R1.

Parasuraman, R. (2003). Neuroergonomics: Research and practice. *Theoretical Issues in Ergonomics Science*, *4*(1-2), 5–20.

Song, Y., Zheng, Q., Liu, B., & Gao, X. (2023). EEG conformer: Convolutional transformer for EEG decoding and visualization. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *31*, 710–719.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *Proceedings of the 31st International Conference on Neural Information Processing Systems*, 6000–6010.

Veličković, P., Cucurull, G., Casanova, A., Romero, A., Liò, P., & Bengio, Y. (2018). Graph attention networks. *International Conference on Learning Representations*.

Wang, F., Wu, S., Zhang, W., Xu, Z., Zhang, Y., & Chu, H. (2020). Multiple nonlinear features fusion based driving fatigue detection. *Biomedical Signal Processing and Control*, *62*, 102075.

Wang, S., Tian, J., Liang, P., Xu, X., Yu, Z., Liu, S., & Zhang, D. (2024). Single and simultaneous fault diagnosis of gearbox via wavelet transform and improved deep residual network under imbalanced data. *Engineering Applications of Artificial Intelligence*, *133*, 108146.

Wickens, C. D. (2008). Multiple resources and mental workload. *Human Factors*, *50*(3), 449–455.

Zhao, C., Zhao, M., Yang, Y., Gao, J., Rao, N., & Lin, P. (2017). The reorganization of human brain networks modulated by driving mental fatigue. *IEEE Journal of Biomedical and Health Informatics*, *21*(3), 743–755. https://doi.org/10.1109/JBHI.2016.2544061

Zheng, W.-L., & Lu, B.-L. (2017). A multimodal approach to estimating vigilance using EEG and forehead EOG. *Journal of Neural Engineering*, *14*, 026017.