Dual-Stage Domain Adaptation with Semantic Center Alignment for Cross-Subject EEG-Based Depression Screening

Ziyue Jiang[1], TIANCHENG CAO[1], Tong Zou[1], Chen Shen[1], and Hen-Wei Huang[1]

[1]Affiliation not available

February 27, 2026

Abstract

Current diagnosis of Major Depressive Disorder (MDD) relies heavily on subjective clinical interviews, lacking objective physiological markers. While EEG-based deep learning holds promise, its clinical deployment is severely hindered by inter-subject variability, where models fail to generalize to unseen patients. To bridge this gap, we propose a dual-stage domain adaptation framework combining ResNet-BiLSTM with Domain Adversarial Neural Networks (DANN) and Semantic Center Alignment (SCA) for robust cross-subject depression detection. We transform 128-channel EEG into topology-preserving maps to retain pathologically relevant spatial connectivity and employ a hybrid backbone to capture spatiotemporal features. Specifically, we integrate a two-level adaptation strategy: DANN aligns global feature distributions across subjects via adversarial training, while SCA enforces class-specific compactness by maintaining dynamic prototype centers in a memory bank, mitigating negative transfer. Extensive experiments on the MODMA dataset demonstrate that our method significantly outperforms subject-dependent baselines (59.18%), achieving 79.09% accuracy with improved stability across diverse subjects. By effectively addressing domain shifts through this dual mechanism, our framework offers a more generalizable screening tool that substantially reduces the reliance on individual calibration, paving the way for reliable computer-aided diagnosis.

Dual-Stage Domain Adaptation with Semantic Center Alignment for Cross-Subject EEG-Based Depression Screening

Ziyue Jiang[1], Tiancheng Cao[*1,2], Tong Zou[1], Chen Shen[1], Hen-Wei Huang[1,3]

Abstract— **Current diagnosis of Major Depressive Disorder (MDD) relies heavily on subjective clinical interviews, lacking objective physiological markers. While EEG-based deep learning holds promise, its clinical deployment is severely hindered by inter-subject variability, where models fail to generalize to unseen patients. To bridge this gap, we propose a dual-stage domain adaptation framework combining ResNet-BiLSTM with Domain Adversarial Neural Networks (DANN) and Semantic Center Alignment (SCA) for robust cross-subject depression detection. We transform 128-channel EEG into topology-preserving maps to retain pathologically relevant spatial connectivity and employ a hybrid backbone to capture spatiotemporal features. Specifically, we integrate a two-level adaptation strategy: DANN aligns global feature distributions across subjects via adversarial training, while SCA enforces class-specific compactness by maintaining dynamic prototype centers in a memory bank, mitigating negative transfer. Extensive experiments on the MODMA dataset demonstrate that our method significantly outperforms subject-dependent baselines (59.18%), achieving 79.09% accuracy with improved stability across diverse subjects. By effectively addressing domain shifts through this dual mechanism, our framework offers a more generalizable screening tool that substantially reduces the reliance on individual calibration, paving the way for reliable computer-aided diagnosis.**

Index Terms—**EEG, MDD, Domain Adaptation, Semantic Center Alignment, Cross-subject Generalization.**

I. INTRODUCTION

Major Depressive Disorder (MDD), traditionally conceptualized as a prevalent psychiatric condition characterized by persistent low mood and cognitive impairment, is conventionally diagnosed and monitored through clinical interviews and self-report scales [1]. However, these traditional diagnostic procedures rely heavily on subjective reporting, which can be susceptible to bias and lacks objective physiological evidence [2]. The integration of automated classification methods utilizing Electroencephalogram (EEG) signals represents a valuable resource for clinicians, offering a quantitative and objective biomarker to assist in the precise diagnosis of depression [3,4]. Recent studies have further underscored the importance of secure and robust end-to-end monitoring systems for precision diagnosis, demonstrating the efficacy of hardware-accelerated deep learning in clinical settings [5].

The rapid advancement of machine learning and deep learning techniques has enabled fully automated depression classification using electroencephalography (EEG) data. A wide range of computational approaches have been proposed, many of which have achieved notable performance on established benchmark datasets. [6]. However, most existing methods focus on subject-dependent classification, where training and testing are performed on the same group of participants, limiting their generalizability to unseen individuals [7,8]. This generalization gap occurs because standard deep learning models tend to overfit subject-specific attributes, such as individual baseline neural power and anatomical variability, rather than capturing the subject-invariant pathological features required for reliable cross-subject diagnosis. The substantial inter-subject variability in EEG patterns, arising from differences in skull thickness, electrode impedance, and individual neural signatures, poses significant challenges for cross-subject depression detection—a critical requirement for real-world clinical deployment [9].

Although the physiological basis for inter-subject variability is well-understood, addressing it computationally remains a formidable task. Conventional deep learning models often treat EEG signals as flat vectors or 2D grids without strictly adhering to the non-Euclidean spatial topology of electrode placement, potentially losing critical spatial information associated with brain connectivity [10,11]. Moreover, standard supervised learning algorithms operate under the assumption that training and testing data share identical statistical distributions. In cross-subject scenarios, this assumption is fundamentally violated due to domain shift, where the feature distribution of a new subject (target domain) deviates significantly from that of the training group (source domain) [12]. Consequently, models trained without explicit distribution alignment mechanisms are prone to overfitting the source subjects and fail to extract subject-invariant features necessary for reliable diagnosis.

To bridge these gaps, this study proposes a novel framework (illustrated in Fig. 1) that integrates topology-preserving input with deep spatiotemporal learning and dual-stage domain adaptation. First, we transform raw EEG signals into topological maps (illustrated in Fig. 2) to preserve the non-Euclidean spatial structural relationships of brain activity. Unlike vector-based methods which flatten the signals and inevitably discard spatial adjacency information, this topological transformation provides a biologically plausible representation [13]. This enables the backbone network to explicitly capture intrinsic connectivity patterns distinctive to depressive states. Second, we design a hybrid neural network backbone combining Convolutional Neural Networks (CNN) and Bidirectional Long Short-Term Memory (BiLSTM) units. This architecture allows the model to simultaneously capture local spatial patterns and long-term temporal dependencies, forming a comprehensive representation of the depressive physiological state.

[1]School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore.
[2]Department of Emergency Medicine, Brigham and Women's Hospital, Boston, USA
[3]Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore
*Corresponding author(tiancheng.cao@ntu.edu.sg)

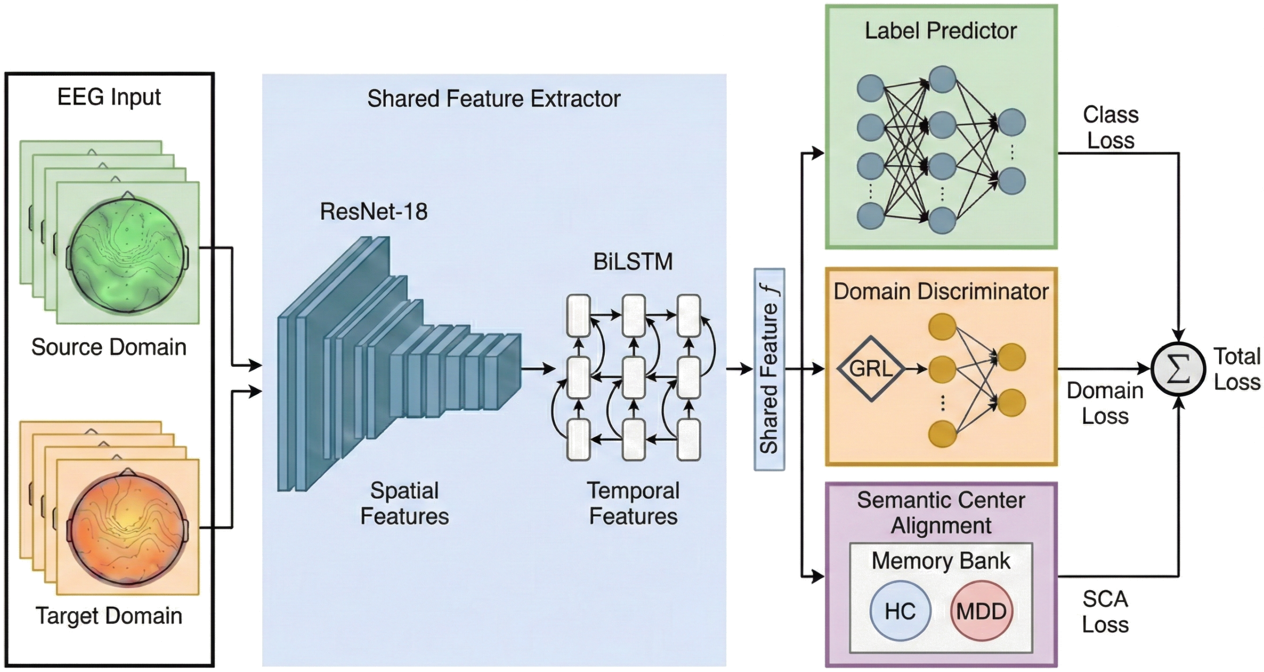

Figure 1. Architecture of the proposed framework. The model comprises a Shared Feature Extractor for spatiotemporal learning, a Depression Classifier for classification, a Domain Discriminator and a Semantic Center Alignment module. A Gradient Reversal Layer (GRL) is used to align feature distributions across different subjects. The SCA module incorporates a dynamic memory bank to enforce semantic consistency. The network is optimized jointly via the weighted sum of classification loss, domain adversarial loss, and SCA loss.

One step further, to enable robust cross-subject generalization, we incorporate a dual-stage domain adaptation strategy. Initially, we employ a Domain Adversarial Neural Network (DANN) to align the global feature distributions across subjects via adversarial training. This approach follows the principle of minimizing the divergence between source and target domains, like strategies employed in Maximum Mean Discrepancy (MMD) [14]. However, standard global alignment methods fundamentally ignore class boundaries. Even recent hierarchical approaches such as Two-level Domain Adaptation (TDANN) [15], which attempt to combine global and local alignment, often rely on statistical matching that lacks explicit semantic constraints. This limitation inherently risks in 'negative transfer,' where samples from different classes are statistically aligned but semantically mismatched. Such semantic misalignment constitutes a critical bottleneck in current domain adaptation methods, limiting their precision on difficult subjects.

To address this limitation, we further introduce a Semantic Center Alignment (SCA) module inspired by [16]. By utilizing a dynamic memory bank to minimize the distance between features and their corresponding class prototypes, SCA explicitly aligns the semantic centers of the source and target domains. This unified framework ensures that the learned features are not only subject-invariant but also semantically discriminative. By effectively bridging the domain gap between subjects, our approach paves the way for a reliable and generalizable screening tool capable of effective assessment in real-world clinical scenarios.

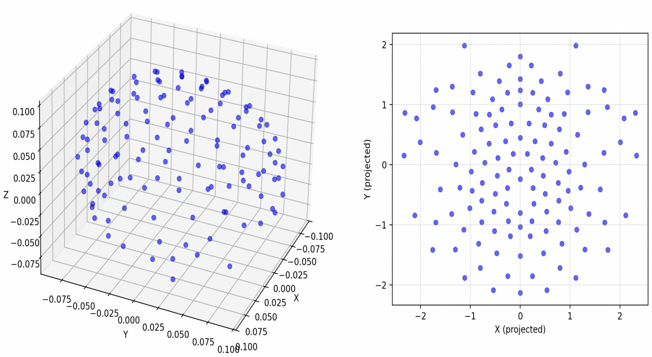

Figure 2. Transformation of EEG spatial topology. The 3D electrode distribution (left) is projected onto a 2D plane (right). This mapping preserves non-Euclidean spatial relationships, enabling the convolutional layers to capture intrinsic brain connectivity.

II. RELATED WORK

Early research in depression detection largely relied on conventional machine learning. As reviewed by Khare *et al.* [6], these methods typically employed handcrafted features—such as spectral power and entropy combined with classifiers like SVM or KNN. While fundamental, these approaches require extensive feature engineering and often fail to capture abstract neural representations.

With the rise of deep learning, end-to-end frameworks have become dominant. For instance, Zheng *et al.* [8] utilized deep neural networks to identify critical frequency bands and channels for affective computing. More recently, Sharma *et al.* [17] applied standard Convolutional Neural Networks (CNN) to automatically learn distinctive features from EEG time-frequency representations. Similarly, recent advances have leveraged Discrete Wavelet Transform (DWT) within efficient architectures like PoolFormer to enhance wearable monitoring performance [18]. Meanwhile, other approaches have integrated Continuous Wavelet Transform (CWT) with hardware-friendly binary CNNs to optimize physiological signal classification on neuromorphic platforms [19]. Mumtaz and Qayyum [20] developed a deep learning framework combining 1D-CNNs with LSTM units to model temporal dependencies in EEG signals for automated depression screening. Despite these advancements, most existing methods remain subject-dependent. They often suffer from performance degradation when applied to unseen subjects, primarily due to the high inter-subject variability of EEG signals [6, 17]. Moreover, for real-world clinical deployment, robust performance must be maintained not only against domain shifts but also against physical constraints, such as parasitic effects in hardware implementations [21].

To address the limitations of subject-dependent models, cross-subject EEG analysis has gained increasing attention. A recent comprehensive review by Iqbal *et al.* [22] highlights that the primary challenge in this setting lies in the non-stationary nature of EEG signals and the high inter-subject variability, which often leads to poor generalization on unseen data. To mitigate this, Domain Adaptation (DA) techniques have been introduced to align the feature distributions of different domains. Adversarial training strategies, as recently advanced by Yang *et al.* [23] in the context of graph-based EEG analysis, have proven effective in learning domain-invariant representations by aligning global feature distributions. However, standard global alignment methods operate without considering class label information in the target domain. This can lead to the misalignment of complex multimode structures, where samples from different classes may be mapped incorrectly in the feature space, a phenomenon known as 'negative transfer'.

Consequently, recent studies in computer vision have proposed semantic alignment approaches, such as moving centroid alignment [18], to enforce class-level consistency. Despite their success in image recognition, such fine-grained adaptation strategies remain underexplored in EEG-based mental health assessment. While recent studies like Li et al. [24] and Rahul et al. [25] have advanced spatiotemporal feature fusion and deep architectures, they lack explicit mechanisms to align class-conditional distributions across subjects. This necessitates a unified framework capable of simultaneously capturing deep spatiotemporal dynamics and enforcing semantic consistency to ensure robust cross-subject generalization, as proposed in this study.

III. Methods

A. Overview of the Proposed Framework

To strictly preserve the non-Euclidean spatial topology of EEG signals while capturing temporal dynamics and mitigating inter-subject variability, we propose a hybrid deep learning framework. As illustrated in Fig. 1, the architecture consists of four integrated modules: (1) a Shared Feature Extractor that learns discriminative spatiotemporal representations; (2) a Depression Classifier for disease diagnosis; (3) a Domain Discriminator for global adversarial alignment; and (4) a Semantic Center Alignment (SCA) module for class-level semantic consistency. The network is optimized jointly via a weighted sum of classification, domain adversarial, and semantic alignment losses to ensure robust cross-subject generalization.

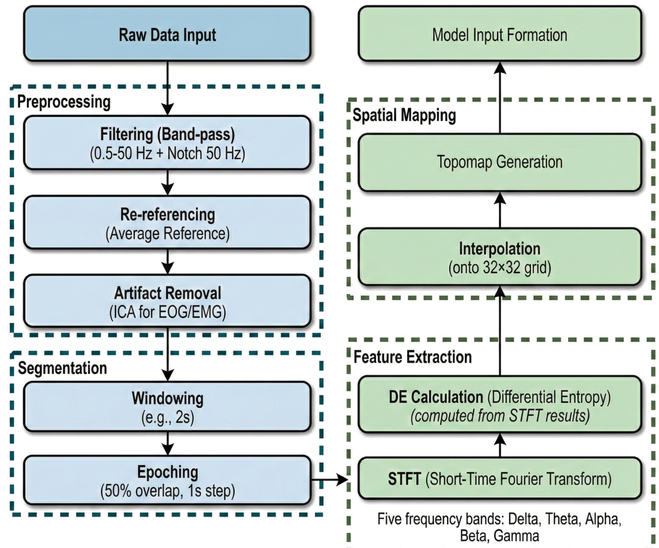

Figure 3. The data processing pipeline. EEG signals undergo preprocessing, segmentation, and DE extraction via STFT to generate 2D topomaps for model input.

B. Spatiotemporal Feature Extraction

The input to our framework consists of topological EEG maps constructed from 128-channel recordings. Following preprocessing and feature extraction (detailed in Section IV.A and Fig. 3), Differential Entropy features from five frequency bands—Delta (1–4 Hz), Theta (4–8 Hz), Alpha (8–13 Hz), Beta (13–30 Hz), and Gamma (30–50 Hz)—are spatially interpolated onto a 32×32 grid based on electrode coordinates. The input EEG sequence is denoted as $X \in R^{B \times T \times C \times H \times W}$, where B is the batch size, T is the number of time frames, and $C = 5$ corresponds to the five frequency bands that comprehensively capture neural oscillatory patterns associated with depression.

To extract spatial representations, we employ ResNet-18 as the backbone network. Given that EEG topological maps contain complex non-Euclidean structural information, the deep convolutional layers of ResNet can capture local spatial patterns and global inter-region connectivity distinctive to depressive brain states. Specifically, the first convolutional layer is adapted to accept 5-channel inputs, matching the spectral dimensions of our topographic maps. A Time-Distributed strategy is applied to reshape the input tensor to $(B \cdot T, 5, H, W)$ allowing the ResNet to process each time frame independently. The output features from the global average pooling layer are flattened into a 512-dimensional vector and a Dropout layer ($p = 0.5$) is applied immediately after the spatial feature extraction to prevent overfitting.

Following spatial encoding, the sequence of feature vectors is fed into a Bidirectional Long Short-Term Memory (BiLSTM) network to model temporal dependencies. Since EEG signals are inherently dynamic time-series data, static spatial features alone are insufficient. The BiLSTM module is explicitly designed to model the temporal evolution and long-range dependencies of brain activity, capturing how depressive states manifest over time. The hidden state size of the LSTM is set to 64. By concatenating the final hidden states from both the forward and backward directions, we obtain a comprehensive global spatiotemporal representation $f \in R^{128}$ (i.e., 64×2), which serves as the input for the subsequent classification tasks.

C. Dual-Stage Domain Adaptation

To enable robust cross-subject generalization, we incorporate a dual-stage domain adaptation strategy that operates in parallel with the main Depression Classifier (G_y). While the classifier ensures discriminative performance on the source domain, the adaptation modules interact with the shared feature f to enforce subject invariance at both global and semantic levels.

1) *Domain Adversarial Neural Networks:* To address global domain shifts, we employ a Domain Adversarial Neural Network (DANN). A Domain Discriminator (G_d) is introduced to predict the subject identity. Significantly, a Gradient Reversal Layer (GRL) is inserted before the discriminator. During forward propagation, the GRL acts as an identity transform; during backpropagation, it reverses the gradient by multiplying it by a negative scalar $-\lambda$. This creates a minimax adversarial game: while G_d attempts to minimize subject classification error, the feature extractor maximizes it, effectively stripping away subject-specific variations.

2) *Semantic Center Alignment*: Standard DANN aligns global distributions but ignores class boundaries, potentially leading to 'negative transfer' where samples from different classes are mapped incorrectly. To mitigate this, we introduce the SCA module, which maintains a dynamic Memory Bank storing global prototype centers for Healthy Control (C_{HC}) and MDD (C_{MDD}) categories. During training, the module minimizes the Euclidean distance between the extracted feature f and its corresponding class prototype. This explicit constraint ensures that the adaptation process respects class boundaries, yielding features that are not only subject-invariant but also semantically discriminative.

D. Optimization Objective

The proposed method framework is trained through a joint optimization process that balances two competing objectives: maximizing depression classification accuracy while minimizing subject-specific feature representations. The overall training objective is formulated as:

$$L = L_t + \lambda_d \cdot L_d + \beta \cdot L_s \quad (1)$$

where L_t, L_d, and L_s denote the classification loss, domain adversarial loss, and semantic alignment loss, respectively. λ_d and β are trade-off hyperparameters balancing the contributions of domain adaptation and semantic consistency.

The depression classification loss L_t is computed using label-smoothed cross-entropy to prevent overfitting:

$$L_t = -\sum_{i=1}^{N} [y_i \log(\hat{y}_i) + (1-y_i)\log(1-\hat{y}_i)] \quad (2)$$

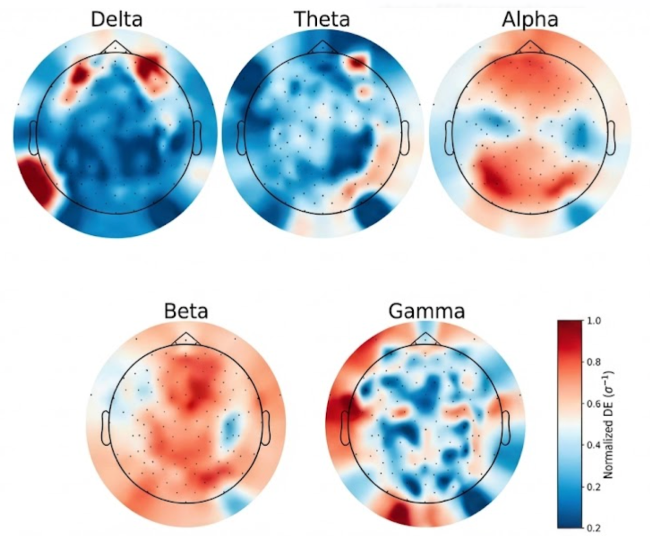

Figure 4. Visualization of Differential Entropy (DE) topomaps. Color intensity represents the spectral energy distribution across brain regions used for depression detection.

where N represents the mini-batch size, and $\hat{y}_i$ is the probability predicted by the depression classifier G_y. The target y_i is the smoothed ground-truth label (with smoothing factor $\varepsilon = 0.1$):

$$y_i \leftarrow (1-\varepsilon)y_i + \frac{\varepsilon}{2} \quad (3)$$

The domain adversarial loss L_d measures the domain discriminator's ability to identify subject identity:

$$L_d = -\frac{1}{N}\sum_{i=1}^{N}\sum_{s=1}^{S} d_{is}\log(\widehat{d_{is}}) \quad (4)$$

where S denotes the total number of subjects in the source domain, and d_{is} is the binary indicator ($d_{is} = 1$ if sample i belongs to subject s, else 0). $\hat{d}_{is}$ represents the predicted subject probability from the domain discriminator G_d.

Finally, the Semantic Center Alignment loss L_s is defined as the squared Euclidean distance between the feature representation and its corresponding class prototype:

$$L_s = \frac{1}{2}\sum_{i=1}^{N} \left\| f_i - C_{y_i} \right\|_2^2 \quad (5)$$

where f_i is the extracted feature vector, and C_{y_i} is the global prototype center corresponding to the ground-truth class y_i, retrieved from the dynamic memory bank. Note that L_s is computed exclusively on source samples as target labels are unavailable.

IV. Experiments

To evaluate the effectiveness of the proposed method framework for cross-subject depression detection, extensive experiments were conducted on the MODMA dataset. This section details the experimental setup, evaluation protocols, and quantitative analysis.

A. Dataset and Data Processing

In this study, we utilize the 128-channel EEG subset of the Multi-modal Open Dataset for Mental-disorder Analysis (MODMA) [26]. The dataset comprises 53 participants, consisting of 24 patients clinically diagnosed with Major Depressive Disorder (MDD) and 29 healthy controls (HC) matched for age and gender. The EEG data used in this study were collected during a resting-state protocol with eyes closed. This condition is widely regarded as stable and effective for analyzing intrinsic brain connectivity patterns associated with depressive disorders, minimizing external visual stimuli interference.

Prior to feature extraction, the raw EEG signals undergo a rigorous preprocessing pipeline as illustrated in Fig. 4. First, a band-pass filter (0.5–50 Hz) and a notch filter (50 Hz) are applied to remove power line interference. Signals are then re-referenced to the average reference, and artifacts (e.g., EOG, EMG) are removed via Independent Component Analysis (ICA). Following segmentation, Differential Entropy (DE) features are extracted from five frequency bands (Delta to Gamma) using Short-Time Fourier Transform (STFT). Finally, to generate the model input X, these DE features are spatially interpolated onto 32×32 grid based on electrode coordinates, preserving the topological structure of brain activity. This generated map is then directly fed into the hybrid neural network for feature learning.

B. Experimental Setup

To comprehensively evaluate the cross-subject generalization capability of the proposed framework in diagnosing unseen patients, we adopted a Leave-One-Subject-Out (LOSO) cross-validation strategy, consistent with established protocols for cross-subject transfer. Specifically, in each iteration, one subject was randomly designated as the test set (target domain), while the remaining subjects constituted the training set (source domain), ensuring that the model is evaluated on strictly unseen patient data to prevent data leakage. Within the training process, 15% of the source data was randomly reserved as a validation set to monitor convergence and select the optimal model.

All methods were implemented in Python using the PyTorch framework on an NVIDIA GPU. We utilized the Adam optimizer with an initial learning rate of 0.0005 and a weight decay of 5×10^{-4} to minimize the objective function. The batch size was set to 32, and the training spanned 40 epochs with a Cosine Annealing scheduler dynamically adjusting the learning rate. For the hyperparameter configuration, the domain adversarial loss weight λ_d and the semantic alignment weight β were both empirically set to 0.1 to balance feature confusion and class compactness, while the GRL adaptation factor λ was scheduled to gradually increase from 0 to 1.0 during training to stabilize the adversarial interplay.

B. Experimental Results

Our evaluation includes a performance comparison to assess cross-subject generalization and an ablation study to verify module effectiveness. Mean accuracy and standard deviation are reported to assess performance stability across subjects. The quantitative results are summarized in Table I and Table II, respectively.

TABLE I. COMPARISON OF CLASSIFICATION PERFORMANCE

Methods	Backbone	Adaptation Strategy	Accuracy (%)
SVM	SVM	None	56.73
DANN [12]	ResNet-BiLSTM	DANN	71.64
MMD[14]	ResNet-BiLSTM	MMD	73.46
TDANN[15]	ResNet-BiLSTM	DANN+ MMD	77.78
Proposed	ResNet-BiLSTM	DANN+SCA	79.09

TABLE II. ABLATION STUDY OF COMPONENT CONTRIBUTIONS

Variants	Global Adaptation	Semantic Alignment	Accuracy (%)
Baseline	✗	✗	59.18
Model A	✓	✗	71.64
Model B	✗	✓	68.24
Proposed	✓	✓	79.09

Comparison of Classification Performance The classification results using the Leave-One-Subject-Out (LOSO) protocol are summarized in Table I. As observed, the baseline ResNet-BiLSTM model showed limited generalization (59.18%) due to significant inter-subject variability, supporting the premise in Section I that domain shift is a primary challenge. Incorporating global domain adaptation (DANN [12], MMD [14]) yielded substantial improvements, demonstrating the necessity of feature distribution alignment.

A key observation is that the comparison with TDANN [15] (77.78%) provides valuable insights into the domain adaptation mechanism. TDANN represents a strong baseline that integrates two effective statistical alignment techniques (global DANN and local MMD). In contrast, even with this combination, its performance remains slightly lower than that of our proposed framework. This suggests that relying solely on statistical distribution matching may face limitations as it does not explicitly account for the semantic consistency of the features. The further improvement of +1.31% achieved by our method (79.09%) indicates that exploring semantic alignment (via SCA) is a promising and feasible direction. It effectively mitigates the "negative transfer" issue that statistical methods alone may not fully resolve.

Ablation Study and Clinical Implications To verify the specific contribution of each component, we conducted an ablation study (Table II). While global adversarial training (Model A) significantly improved performance to 71.64%, the full framework integrating the SCA module achieved the highest accuracy of 79.09%. This incremental improvement of +7.45% over the global-only model confirms the benefit of aligning class-specific semantic centers for learning discriminative subject-invariant features.

From a clinical perspective, the most valuable aspect of these results lies in the experimental setting itself. All reported accuracies were achieved under a strict LOSO protocol, meaning the model was tested on completely unseen patients. The high accuracy in this rigorous scenario validates the framework's strong generalization capability. Unlike subject-dependent models, our method demonstrates consistent performance across diverse subjects, confirming its potential as an objective and reliable auxiliary tool for large-scale community health screening.

V. Conclusion

This work addresses a fundamental challenge in EEG-based mental health screening: poor generalization to unseen patients due to inter-subject variability. We proposed a dual-stage domain adaptation framework where Domain Adversarial Neural Networks (DANN) align global feature distributions, while Semantic Center Alignment (SCA) enforces class-specific structure through dynamic prototype learning. By integrating these adaptation mechanisms with topology-preserving EEG representations that capture intrinsic spatial connectivity, our approach explicitly prevents negative transfer and ensures semantic consistency. Experimental results on the MODMA dataset demonstrate a superior accuracy of 79.09% under a strict Leave-One-Subject-Out protocol. This confirms that our framework achieves robust cross-subject generalization without individual calibration, demonstrating significant potential for scalable and objective community health screening. Future directions include multi-site validation and attention-based interpretability for biomarker discovery.

References

[1] R. Uher, J. L. Payne, B. Pavlova, and R. H. Perlis, "Major depressive disorder in DSM-5: Implications for clinical practice and research of changes from DSM-IV," *Depression and Anxiety*, vol. 31, no. 6, Jun. 2014, pp. 459-471.

[2] D. A. Regier, W. E. Narrow, D. E. Clarke, H. C. Kraemer, S. J. Kuramoto, E. A. Kuhl, and D. J. Kupfer, "DSM-5 field trials in the United States and Canada, Part II: Test-retest reliability of selected categorical diagnoses," *American Journal of Psychiatry*, vol. 170, no. 1, Jan. 2013, pp. 59-70.

[3] S. Olbrich and M. Arns, "EEG biomarkers in major depressive disorder: Discriminative power and prediction of treatment response," *International Review of Psychiatry*, vol. 25, no. 5, Oct. 2013, pp. 604-618.

[4] A. Sun, J. Sun, X. Chen, and X. Gao, "A data-centric and interpretable EEG framework for depression severity grading using SHAP-based insights," *Journal of NeuroEngineering and Rehabilitation*, vol. 22, no. 1, 2025, pp. 1-16.

[5] T. Cao, W. S. Ng, C. Shen, H. Li, R. Tan, D. Wang, and H.-W. Huang, "Cybersecure End-to-End FPGA-Accelerated ECG Monitoring for Precision Diagnosis With Personalized CWT and Adversarial Defense," *IEEE Journal of Biomedical and Health Informatics*, vol. 29, no. 12, pp. 8863-8870, 2024.

[6] S. K. Khare, V. Bajaj, and U. R. Acharya, "Depression detection and diagnosis based on electroencephalogram (EEG) analysis: A systematic review," *Brain Sciences*, vol. 14, no. 11, Nov. 2024, Art. no. 1107.

[7] Z. Zhang *et al.*, "TSF-MDD: A deep learning approach for electroencephalography-based diagnosis of major depressive disorder with temporal–spatial–frequency feature fusion," *Bioengineering*, vol. 12, no. 1, Jan. 2025, Art. no. 85.

[8] W.-L. Zheng and B.-L. Lu, "Investigating critical frequency bands and channels for EEG-based emotion recognition," *IEEE Transactions on Autonomous Mental Development*, vol. 7, no. 3, Sep. 2015, pp. 162-175.

[9] F. Lotte, L. Bougrain, A. Cichocki, M. Clerc, M. Congedo, A. Rakotomamonjy, and F. Yger, "A review of classification algorithms for EEG-based Brain–Computer interfaces: A 10-year update," *Journal of Neural Engineering*, vol. 15, no. 3, Art. no. 031005, 2018.

[10] U. R. Acharya, S. L. Oh, Y. Hagiwara, J. H. Tan, and H. Adeli, "Automated EEG-based screening of depression using deep convolutional neural network," *Computer Methods and Programs in Biomedicine*, vol. 161, pp. 103-113, 2018.

[11] T. Song, W. Zheng, P. Song, and Z. Cui, "EEG emotion recognition using dynamical graph convolutional neural networks," *IEEE Transactions on Affective Computing*, vol. 11, no. 3, pp. 532-541, Jul.-Sep. 2020.

[12] Y. Ganin et al., "Domain-adversarial training of neural networks," *Journal of Machine Learning Research*, vol. 17, no. 1, pp. 2096-2030, 2016.

[13] P. Bashivan, I. Rish, M. Yeasin, and N. Codella, "Learning representations from EEG with deep recurrent-convolutional neural networks," in International Conference on Learning Representations (ICLR), 2016.

[14] A. Gretton, K. M. Borgwardt, M. J. Rasch, B. Schölkopf, and A. Smola, "A kernel two-sample test," *Journal of Machine Learning Research*, vol. 13, pp. 723-773, Mar. 2012.

[15] G. Bao et al., "Two-Level Domain Adaptation Neural Network for EEG-Based Emotion Recognition," *Frontiers in Human Neuroscience*, vol. 14, p. 605246, Jan. 2021.

[16] S. Xie, Z. Zheng, L. Chen, and C. Chen, "Learning Semantic Representations for Unsupervised Domain Adaptation," in *Proceedings of the International Conference on Machine Learning*, 2018, pp. 5423-5432.

[17] G. Sharma, A. Parashar, and A. M. Joshi, "DepNet: A deep convolutional neural network framework for detecting depression using EEG signals," *IEEE Sensors Journal*, vol. 21, no. 23, pp. 26346-26357, Dec. 2021.

[18] T. Cao, W. S. Ng, W. L. Goh, and Y. Gao, "DWT-PoolFormer: Discrete Wavelet Transform-based Quantized Parallel PoolFormer Network Implemented in FPGA for Wearable ECG Monitoring," in 2024 *IEEE Biomedical Circuits and Systems Conference (BioCAS)*, 2024, pp. 1–5.

[19] T. Cao, Z. Zhang, W. L. Goh, C. Liu, Y. Zhu, and Y. Gao, "ECG classification using binary CNN on RRAM crossbar with nonidealities-aware training, readout compensation and CWT preprocessing," in 2023 *IEEE Biomedical Circuits and Systems Conference (BioCAS)*, 2023, pp. 1–5.

[20] W. Mumtaz and A. Qayyum, "A deep learning framework for automatic diagnosis of unipolar depression," International Journal of Medical Informatics, vol. 132, p. 103983, Dec. 2019.

[21] T. Cao, C. Liu, Y. Gao, and W. L. Goh, "Parasitic-aware modelling for neural networks implemented with memristor crossbar array," in 2021 *IEEE 14th International Symposium on Embedded Multicore/Many-core Systems-on-Chip (MCSoC)*, 2021, pp. 1–7.

[22] M. Z. Iqbal et al., "Opportunities and Challenges for Clinical Practice in Detecting Depression Using EEG and Machine Learning," *Sensors*, vol. 25, no. 2, Art. no. 409, Jan. 2025.

[23] J. Yang et al., "Major Depressive Disorder Detection Using Graph Domain Adaptation with Global Message-Passing Based on EEG Signals," *IEEE Transactions on Emerging Topics in Computing, early access*, pp. 1-14, 2025.

[24] Y. Li, X. Zhang, and J. Wang, "Graph-based EEG approach for depression prediction: integrating time-frequency complexity and spatial topology," *Frontiers in Neuroscience*, vol. 18, Art. no. 1367212, Apr. 2024.

[25] S. A. Rahul, K. Kumar, and M. Singh, "Decentralized EEG-based detection of major depressive disorder via transformer architectures and split learning," *Frontiers in Computational Neuroscience*, vol. 19, Art. no. 1569828, Feb. 2025.

[26] H. Cai et al., "A multi-modal open dataset for mental-disorder analysis," *Scientific Data*, vol. 9, no. 1, Art. no. 178, 2022.