PR-DA: Prototype Regularization Domain Adaptation for Cross-Subject EEG-Based Emotion Recognition

1st Rongtao Chen
School of artificial intelligence
South China Normal University
Foshan, Guangdong Province, China
chenrongtao@m.scnu.edu.cn

1st Zhepei Hong
School of artificial intelligence
South China Normal University
Foshan, Guangdong Province, China
hongzhepei@m.scnu.edu.cn

3rd Qi You
School of artificial intelligence
South China Normal University
Foshan, Guangdong Province, China
youqi@m.scnu.edu.cn

4th Chuwen Xie
School of artificial intelligence
South China Normal University
Foshan, Guangdong Province, China
xcwxcw@m.scnu.edu.cn

5th Jiahui Zhang
School of artificial intelligence
South China Normal University
Foshan, Guangdong Province, China
zjhzjh@m.scnu.edu.cn

6th Jiahui Pan
School of artificial intelligence
South China Normal University
Foshan, Guangdong Province, China
panjiahui@m.scnu.edu.cn

***Abstract*—Electroencephalogram (EEG)-based emotion recognition holds significant potential in healthcare, traffic safety, and entertainment. However, cross-subject emotion recognition remains challenging due to individual differences and the difficulty in extracting domain-invariant features. To address these issues, this paper proposes a novel Prototype Regularization Domain Adaptation (PR-DA) framework. Experimental results on three benchmark datasets (SEED, SEED-IV, and SEED-VII) demonstrate that the proposed PR-DA framework achieves superior performance compared to state-of-the-art methods, with accuracies of 96.30%±2.87%, 86.56%±4.67%, and 50.43%±6.71%, respectively. The proposed PR-DA framework offers a promising approach for cross-subject EEG-based emotion recognition. The source code is available at the following link: https://github.com/seizeall/PR-DA.**

***Index Terms*—Electroencephalogram (EEG), emotion recognition, cross-subject, prototype regularization, domain adaptation.**

I. Introduction

Emotion recognition has garnered significant interest due to its wide-ranging applications in healthcare, traffic safety, and entertainment [1]. However, non-physiological signals are less reliable, as they are susceptible to subjective manipulation. In contrast, electroencephalography (EEG) offers a more robust approach by directly measuring complex neural activity and brain connectivity patterns associated with emotional states.

Despite its potential, EEG-based emotion recognition faces a critical challenge in cross-subject scenarios due to individual differences and the difficulty in extracting domain-invariant features. Recent studies indicate that cross-subject EEG-based emotion recognition consistently underperforms compared to within-subject methods. For instance, Zhu et al. [2] utilized a linear support vector machine (SVM) for emotion recognition on the SEED dataset [3], achieving an average accuracy of 64.82% for cross-subject recognition, significantly lower than the 90.97% for within-subject recognition. Similarly, Nath et al. [4] employed long short-term memory networks and linear SVM on the SEED dataset, reporting an average accuracy of 71.22% for cross-subject models, compared to 93.91% for within-subject models. Thus, improving the accuracy of cross-subject emotion recognition remains a critical challenge in this field, essential for advancing brain-computer interface applications. The challenges in cross-subject emotion recognition can be summarized as follows:

1) **Extraction of Domain-Invariant Features:** The extraction of domain-invariant features with salient emotional information remains a primary challenge. Graph structures can naturally represent the connectivity relationships between brain regions and have demonstrated high efficiency in analyzing neural mechanisms related to emotions. Consequently, graph convolutional networks (GCN) have become an effective technique for processing EEG signals [5]. Numerous studies have recently focused on exploring the potential of GCN in extracting domain-invariant features. For example, Song et al. [6] proposed DGCNN, which dynamically constructs an adjacency matrix during training to effectively capture both spatial and functional connectivity among EEG channels. Jin et al. [7] developed PGCN, which aggregates features across the entire graph to learn relationships between EEG electrodes, thereby enhancing the representation of EEG signals for emotion recognition. However, traditional GCN models primarily focus on modeling spatial graph representations, which fail to fully capture the complexity and domain-invariant features of brain structure during dynamic mood changes.

2) **Alignment of Joint Probability Distributions:** Dynamically aligning the joint probability distributions of different

Corresponding authors: Qi You and Jiahui Pan

individuals remains a major obstacle. EEG signals exhibit significant inter-individual variability, making it challenging to align joint probability distributions in the same feature space [8]. Recent research has increasingly focused on domain adaptation (DA) frameworks to mitigate individual differences. For instance, Zhou et al. [9] proposed PR-PL, which combines prototypical representation and pairwise learning. This method creates prototypes for each emotion and uses pairwise loss to enhance intra-class compactness and inter-class separability. Guo et al. [10] presented MSDA-SFE, which employs a shared feature extractor across multiple source domains to learn domain-invariant features and adapts them to the target domain via domain-specific classifiers. Despite these achievements, current domain adaptation frameworks typically focus on capturing domain differences to update models. This approach emphasizes marginal probability distributions while neglecting conditional probability distributions during data alignment.

To address these challenges, we propose PR-DA, a prototype regularization domain adaptation framework for cross-subjec EEG-based emotion recognition. The key contributions of this work are threefold:

- We propose a Connection Uncertainty Spatial-Temporal Graph (CUSTG) module as a shared feature extractor for the DA architecture, which dynamically constructs adaptive graph representations and extracts domain-invariant neural signatures using spatio-temporal features.
- We introduce a prototype regularization strategy that enhances domain adaptation. This method combines a prototype contrastive loss to align features with class prototypes, improving intra-class compactness and inter-class separability.
- We perform extensive cross-subject experiments on three benchmark emotion datasets: SEED, SEED-IV, and SEED-VII. Results show our PR-DA framework outperforms state-of-the-art methods.

II. Method

We assume that the EEG samples in the source and target domains are denoted as $D_s = \{X_s, Y_s\}$ and $D_t = \{X_t, Y_t\}$, respectively. Here, $\{X_s, Y_s\} = \{(x_i^s, y_i^s)\}_{i=1}^{N_s}$ and $\{X_t, Y_t\} = \{(x_i^t, y_i^t)\}_{i=1}^{N_t}$, where x_i^s and x_i^t are the EEG samples, and y_i^s and y_i^t are the corresponding emotion labels. Note that y_i^t is not available during the model training process. The complete architecture of our proposed model is shown in Fig. 1. To further elucidate the model's structure, we present the detailed structure of the CUSTG Module in Fig. 2.

A. Connection Uncertainty Spatial-Temporal Graph

The primary objective of CUSTG is to explicitly model the uncertainty of neural connections by leveraging shared temporal features and multi-scale spatial features, thereby extracting more robust and domain-invariant spatio-temporal transition features for subsequent prototype regularization tasks.

1) Temporal Dependency Modeling: To capture the time-varying relationships between EEG channels, we first model the long-range temporal dependencies within each channel's signal. Given an input EEG feature tensor $\boldsymbol{X} \in \mathbb{R}^{B \times N_c \times D_{in}}$, where B is the batch size, N_c is the number of channels, and D_{in} is the input feature dimension, we first process it with a Bidirectional Gated Recurrent Unit (BiGRU) to obtain deep temporal representations:

$$\boldsymbol{H}^{temp} = \text{BiGRU}(\boldsymbol{X}), \tag{1}$$

where $\boldsymbol{H}^{temp}$ represents the concatenated hidden states from both forward and backward directions. Crucially, $\boldsymbol{H}^{temp}$ is shared across all scales of spatial feature extraction to provide a consistent temporal context.

To dynamically estimate the connection probabilities, we employ a temporal edge predictor. This module, composed of a two-layer MLP with a ReLU activation, takes pairs of temporal node features from $\boldsymbol{H}^{temp}$ to quantify their connection potential. For any two node features $(\boldsymbol{h}_i, \boldsymbol{h}_j)$, it computes a logit vector $\boldsymbol{l}_{ij} \in \mathbb{R}^2$:

$$\boldsymbol{l}_{ij} = \boldsymbol{W}_2 \cdot \text{ReLU}(\boldsymbol{W}_1 \cdot [\boldsymbol{h}_i || \boldsymbol{h}_j] + \boldsymbol{b}_1) + \boldsymbol{b}_2, \tag{2}$$

where $[\boldsymbol{h}_i || \boldsymbol{h}_j]$ is the concatenation of the feature vectors. Here, $\boldsymbol{W}_1$ and $\boldsymbol{W}_2$ are the learnable weight matrices.

To generate a probabilistic adjacency mask while allowing for discrete edge selection during inference, we employ the Gumbel-Softmax trick on the computed logits. This produces a temporal mask matrix $\boldsymbol{P}_t \in [0,1]^{B \times N_c \times N_c}$:

$$\boldsymbol{P}_t[i,j] = \frac{\exp((\boldsymbol{l}_{ij}[1] + g_{ij,1})/\tau)}{\sum_{k=0}^{1} \exp((\boldsymbol{l}_{ij}[k] + g_{ij,k})/\tau)}, \tag{3}$$

where $\boldsymbol{l}_{ij}[1]$ is the logit for an edge existing between nodes i and j. The terms $g_{ij,k} \sim \text{Gumbel}(0,1)$ are independent and identically distributed samples from a Gumbel distribution, which introduce stochasticity during training to explore different graph structures.

The sampling process is carefully controlled throughout training by gradually annealing the temperature parameter τ. For a given training epoch e out of total epochs E, the temperature is updated as:

$$\tau^{(e)} = \max\left(\tau_{end}, \tau_{start} - (\tau_{start} - \tau_{end}) \cdot \frac{e}{E \cdot \eta_a}\right), \tag{4}$$

where η_a is the fraction of epochs over which annealing occurs. Towards the end of training, a hard selection mechanism is enabled to finalize the graph structure. To maintain differentiability for backpropagation, the gradient is approximated using the original soft probabilities via the straight-through estimator:

$$\boldsymbol{P}_t = \text{stop\_gradient}(\hat{\boldsymbol{P}}_{t,\text{hard}}[i,j] - \boldsymbol{P}_t[i,j]) + \boldsymbol{P}_t[i,j], \tag{5}$$

where $\hat{\boldsymbol{P}}_{t,\text{hard}}$ is the one-hot vector derived from the argmax operation and stop_gradient prevents the gradient from flowing through the discrete selection part. The resulting matrix $\boldsymbol{P}_t$ is a globally shared mask that encodes the learned temporal connectivity patterns.

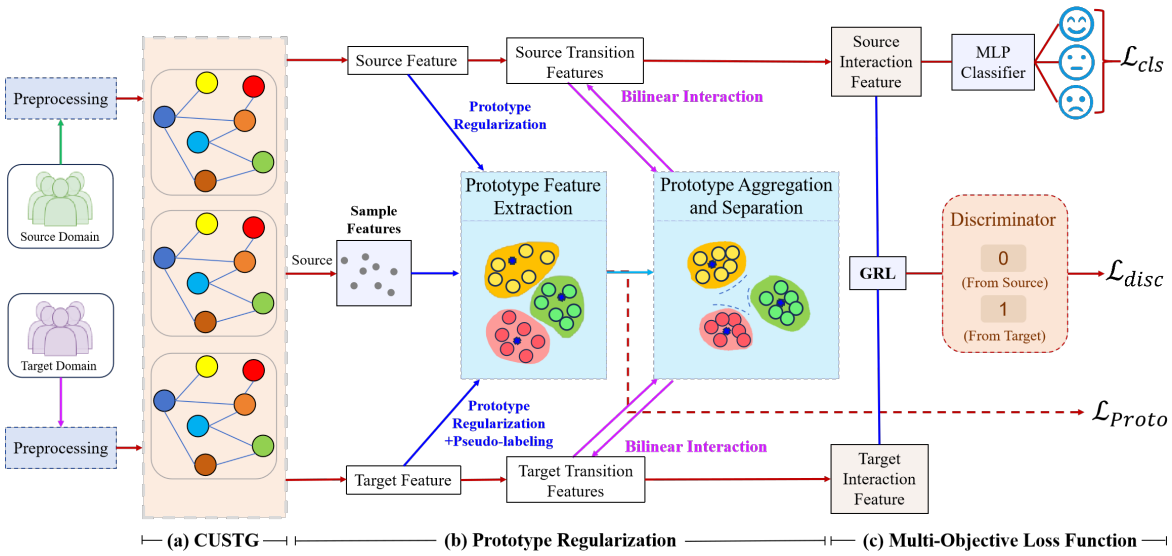

Fig. 1. Framework of the Proposed Prototype Regularization Domain Adaptation.

2) Spatial Feature Modeling: To capture spatial information, we employ a series of convolutional blocks. Each block $i \in \{1, 2, 3\}$ processes the output of the previous one, extracting progressively more abstract spatial features:

$$\boldsymbol{X}_i^{conv} = \text{MaxPool}(\text{Dropout}(\text{ReLU}(\text{Conv2D}_i(\boldsymbol{X}_{i-1}^{conv})))), \tag{6}$$

where $\boldsymbol{X}_0^{conv}$ is the initial input. The output of each block is flattened channel-wise to produce a multi-scale spatial feature map $\boldsymbol{X}_i^{spa}$. For each scale i, we construct a scale-specific graph. The spatial features $\boldsymbol{X}_i^{spa}$ are projected into a bottleneck space:

$$\boldsymbol{H}_i^{spa} = \tanh(\boldsymbol{W}_{bn,i}^{spa} \boldsymbol{X}_i^{spa} + \boldsymbol{b}_{bn,i}^{spa}). \tag{7}$$

Then, a dedicated Spatial Edge Predictor (instantiating the architecture from Eq. 2 with its own weights) is used to generate a spatial mask matrix $\boldsymbol{P}_{s,i}$ from $\boldsymbol{H}_i^{spa}$.

For each scale i, a fused bottleneck representation $\boldsymbol{H}_i^{fused}$ is created by combining the scale-specific spatial features $\boldsymbol{H}_i^{spa}$ and the shared temporal features $\boldsymbol{H}^{temp}$. A primitive affinity matrix $\hat{\boldsymbol{A}}_{prim,i}$ is computed and normalized:

$$\begin{cases} \boldsymbol{H}_i^{fused} & = \tanh(\boldsymbol{W}_{fuse,i}[\boldsymbol{H}_i^{spa} || \boldsymbol{H}^{temp}] + \boldsymbol{b}_{fuse,i}) \\ \hat{\boldsymbol{A}}_{prim,i} & = \text{Softmax}_{\text{row}}(\boldsymbol{H}_i^{fused} \cdot (\boldsymbol{H}_i^{fused})^\top). \end{cases} \tag{8}$$

Final, adaptive spatio-temporal adjacency matrix for scale i, denoted $\boldsymbol{A}_{st,i}$, is formulated as the element-wise product of the three components:

$$\boldsymbol{A}_{st,i} = \hat{\boldsymbol{A}}_{prim,i} \odot \boldsymbol{P}_{s,i} \odot \boldsymbol{P}_t, \tag{9}$$

where $\odot$ denotes the Hadamard product. This triple-gating mechanism ensures that a connection at a given scale is supported by multi-scale feature similarity, scale-specific spatial plausibility, and global temporal coherence.

3) Graph Convolution Module: With the adaptive adjacency matrix $\boldsymbol{A}_{st,i}$ and the corresponding spatial features $\boldsymbol{X}_i^{spa}$ for each scale, we perform graph convolution to aggregate information and learn high-level node representations. As our framework utilizes a dense adjacency matrix, we employ a DenseGCNConv layer for this purpose. For each scale i, the output features $\boldsymbol{F}_i^{trans}$ are computed as:

$$\boldsymbol{F}_i^{trans} = \text{ReLU}(\text{DenseGCNConv}(\boldsymbol{X}_i^{spa}, \boldsymbol{A}_{st,i})), \tag{10}$$

this operation allows each EEG channel to update its representation by integrating information from other channels, weighted by the learned spatio-temporal connectivity strengths.

Finally, the features propagated through the different scale-specific graphs are concatenated to form the final, comprehensive feature representation. This feature vector encapsulates rich spatio-temporal dynamics of the brain:

$$\boldsymbol{F}^{trans} = [\boldsymbol{F}_1^{trans} || \boldsymbol{F}_2^{trans} || \boldsymbol{F}_3^{trans}], \tag{11}$$

this concatenated feature is then flattened and passed to the subsequent modules of the network for classification and domain adaptation.

B. Prototype Regularization

Prototype regularization utilizes domain-invariant transition features extracted by CUSTG to learn representative prototypes for each emotion category. This constructs a feature embedding space that promotes intra-class compactness and inter-class separability among prototypes.

1) Prototype Contrastive Loss: The core of our regularization strategy is the prototype contrastive loss, which optimizes the structure of the base feature space. The comprehensive feature representation $\boldsymbol{F}^{trans}$ is first flattened to produce the base feature vectors $\boldsymbol{f}^{base}$. Within each training iteration, we compute a set of source prototypes, $\boldsymbol{P}_s = \{\boldsymbol{p}_c\}_{c=1}^{N_{cls}} \in$

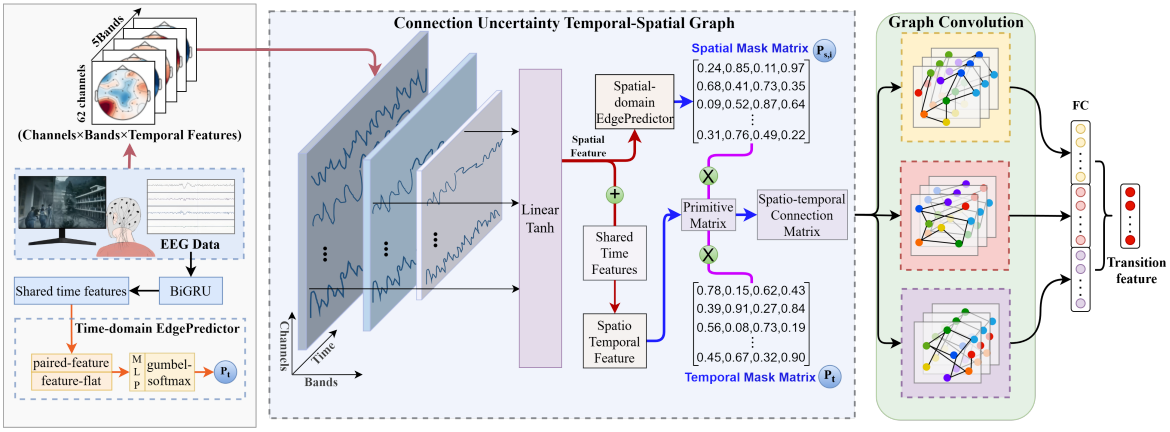

Fig. 2. Detailed Architecture of the Connection Uncertainty Spatial-Temporal Graph Module.

$\mathbb{R}^{N_{cls} \times D_f}$, by averaging the base feature vectors of all source domain samples belonging to class c:

$$\boldsymbol{p}_c = \frac{\sum_{i=1}^{B_s} \mathbb{I}(y_i^s = c) \cdot \boldsymbol{f}_i^{s,base}}{\sum_{i=1}^{B_s} \mathbb{I}(y_i^s = c)}, \tag{12}$$

where y_i^s is the label of the i-th source sample. These prototypes serve as anchors in the feature space.

The contrastive loss then pulls the features of each sample towards their corresponding class prototype while pushing them away from others. For the labeled source domain, the loss is:

$$\mathcal{L}_{S\to P}(\boldsymbol{F}^{s,base}, \boldsymbol{Y}^s, \boldsymbol{P}_s) = -\frac{1}{B_s}\sum_{i=1}^{B_s} \log \frac{\exp(\text{sim}(\boldsymbol{f}_i^{s,base}, \boldsymbol{p}_{y_i^s})/\tau_{cl})}{\sum_{k=1}^{N_{cls}} \exp(\text{sim}(\boldsymbol{f}_i^{s,base}, \boldsymbol{p}_k)/\tau_{cl})}, \tag{13}$$

where $\text{sim}(\cdot,\cdot)$ is the cosine similarity and τ_{cl} is a temperature hyperparameter.

For the unlabeled target domain, we extend the contrastive learning framework by generating online pseudo-labels. Specifically, the pseudo-label for a target sample is determined by assigning the class that corresponds to the highest prediction probability from the model's output. Ultimately, the loss for the target domain samples is formulated over the set:

$$\mathcal{L}_{T\to P}(\boldsymbol{F}^{t,base}, \hat{\boldsymbol{Y}}^t, \boldsymbol{P}_s) = -\frac{1}{B_t'}\sum_{i=1}^{B_t} \mathbb{I}_i \cdot \log \frac{\exp(\text{sim}(\boldsymbol{f}_i^{t,base}, \boldsymbol{p}_{\hat{y}_i^t})/\tau_{cl})}{\sum_{k=1}^{N_{cls}} \exp(\text{sim}(\boldsymbol{f}_i^{t,base}, \boldsymbol{p}_k)/\tau_{cl})}, \tag{14}$$

To further enhance the discriminability of the embedding space, we introduce a prototype separation loss. The prototype separation loss is defined as the average cosine similarity between all pairs of distinct class prototypes:

$$\mathcal{L}_{sep}(\boldsymbol{P}_s) = \frac{1}{N_{cls}(N_{cls}-1)} \sum_{i \neq j} \text{sim}(\boldsymbol{p}_i, \boldsymbol{p}_j), \tag{15}$$

where N_{cls} is the total number of emotion classes.

The total prototype contrastive loss is a weighted sum of the sample-to-prototype contrastive losses for both source and target domains, and the prototype separation loss. The final objective for prototype regularization is given by:

$$\mathcal{L}_{proto} = \frac{1}{2}(\mathcal{L}_{S\to P} + \mathcal{L}_{T\to P}) + \lambda_{ps}\mathcal{L}_{sep}, \tag{16}$$

where λ_{ps} is a hyperparameter that controls the contribution of the prototype separation term.

2) Bilinear Interaction: After structuring the embedding space, we further enhance the features by infusing them with class-specific context through a bilinear interaction mechanism. For each base feature vector $\boldsymbol{f}^{base}$, its interaction with the corresponding class prototype $\boldsymbol{p}_c$ is modeled as:

$$\text{Interaction}(\boldsymbol{f}^{base}, \boldsymbol{p}_c) = \boldsymbol{W}_{inter}(\boldsymbol{f}^{base} \odot \boldsymbol{p}_c), \tag{17}$$

where $\odot$ is the Hadamard product and $\boldsymbol{W}_{inter}$ is a learnable projection matrix. This effect is integrated back into the base feature with a dynamic weight w_b to produce the final, enhanced feature $\boldsymbol{f}^{inter}$:

$$\boldsymbol{f}^{inter} = \boldsymbol{f}^{base} + w_b \cdot \text{Interaction}(\boldsymbol{f}^{base}, \boldsymbol{p}_c), \tag{18}$$

this sequential approach ensures that the features being aligned and classified have already been regularized for maximum class separability.

C. Multi-Objective Loss Function

To effectively train our model, we formulate a multi-objective loss function that simultaneously optimizes for several goals.

1) Classification Loss: The classification loss ensures that the features extracted from the source domain data are discriminative enough to be correctly classified:

$$\mathcal{L}_{cls} = -\frac{1}{N_s}\sum_{i=1}^{N_s}\sum_{k=1}^{C} y_{i,k} \log(p_{i,k}), \tag{19}$$

where N_s is the number of labeled samples in the source domain, C is the number of classes, $y_{i,k}$ is the ground-truth label, and $p_{i,k}$ is the predicted probability of sample i belonging to class k.

2) Domain Discrimination Loss: To encourage the model to learn domain-invariant features, we adoption an adversarial domain discrimination loss. The loss for the domain classifier is typically a binary cross-entropy loss:

$$\mathcal{L}_{dom} = -\mathbb{E}_{\boldsymbol{x}_s \sim \mathcal{D}_s}[\log D(G(\boldsymbol{x}_s))] - \mathbb{E}_{\boldsymbol{x}_t \sim \mathcal{D}_t}[\log(1 - D(G(\boldsymbol{x}_t)))], \tag{20}$$

where $G(\cdot)$ is the feature extractor and $D(\cdot)$ is the domain classifier.

3) Total Loss: The final objective function combines the classification loss, the domain discrimination loss, and the prototype contrastive loss:

$$\mathcal{L}_{total} = \mathcal{L}_{cls} + \lambda_{dom}\mathcal{L}_{dom} + \lambda_{proto}\mathcal{L}_{proto}, \tag{21}$$

where λ_{dom} and λ_{proto} are trade-off hyperparameters that balance the influence of the domain alignment and prototype-based regularization terms, respectively.

III. Experiments

A. Emotion Datasets

The experiments were conducted on three widely used benchmark emotion datasets:

- **SEED [3]** comprises EEG data from 15 subjects across three emotional categories: negative, neutral, and positive. Emotions were elicited using 15 carefully selected 4-minute video clips. Data collection involved three separate sessions per subject on different dates, with each session consisting of 15 EEG recording trials.
- **SEED-IV [11]** includes 15 subjects across three sessions. It employs 24 culturally relevant film clips to induce four emotional states—neutral, sad, fear, and happy—with six unique audiovisual stimuli representing each category.
- **SEED-VII [12]** involves 20 subjects across four sessions. Using 80 video clips, it evokes seven emotions: disgust, fear, sad, neutral, happy, anger, and surprise. EEG data was collected in 20 trials per session, with subjects providing continuous emotional intensity labels.

B. Experiment Results and Analysis

To validate the generalization ability of PR-DA for new subjects, experiments were conducted using the LOSO cross-validation paradigm. Table I presents the prediction results of PR-DA on the SEED and SEED-IV datasets, while Table II shows the prediction results on the SEED-VII dataset. Compared to recent state-of-the-art GCN methods (e.g., MSFR-GCN [13] and CU-GCN [14]) and domain adaptation frameworks (e.g., CL-DDA [15] and DANN-PRLI [16]), PR-DA demonstrates superior predictive performance, achieving the best results. Specifically, PR-DA improves accuracy by 1.85% on the SEED dataset, 0.73% on the SEED-IV dataset, and 1.30% on the SEED-VII dataset compared to various state-of-the-art methods. Furthermore, to provide a more comprehensive evaluation of PR-DA's performance, the AUC (area under the curve) curves for the model's training and testing processes are presented in Fig. 3.

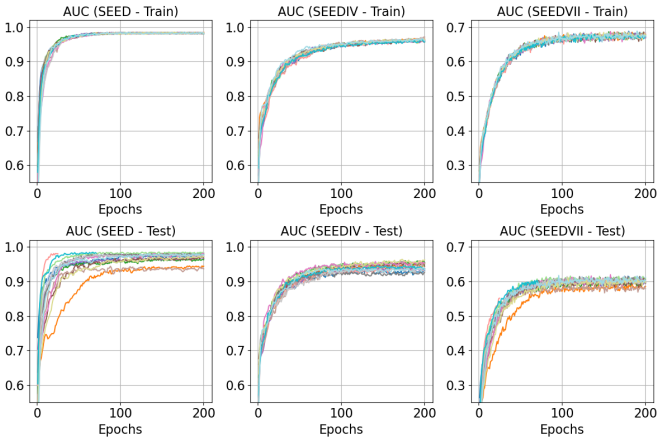

Fig. 3. AUC Curves for the Training and Testing Phases of the Proposed Model on Different Datasets. The lines of different colors represent the results of the experiment for different subjects.

TABLE I
Performance Comparison of PR-DA and State-of-the-Art Methods on SEED and SEED-IV Datasets Using Leave-One-Subject-Out (LOSO) Cross-Validation.

Methods	SEED		SEED-IV	
	P_{acc}	F1 score	P_{acc}	F1 score
MSFR-GCN [13]	86.78±05.40	88.67±06.24	73.43±07.32	73.86±07.52
PGCN [7]	84.59±08.68	84.71±08.64	73.69±07.16	73.75±07.24
CU-GCN [14]	87.10±05.44	87.22±05.49	74.50±07.88	74.64±07.93
BAFNet [17]	93.93±02.76	–	85.83±04.59	–
PR-PL [9]	93.06±05.12	93.12±05.18	81.32±08.53	81.47±08.55
ASJDA [18]	93.75±11.39	93.84±11.36	80.29±19.44	80.37±19.48
S^2A^2-MSDA [19]	90.11±07.32	90.18±07.37	76.23±09.02	76.32±09.09
CL-DDA [15]	90.08±06.04	90.83±06.42	77.55±11.95	77.92±10.87
CMTCF [20]	93.80±04.96	–	79.37±06.05	–
DANN-PRLI [16]	94.45±04.67	94.57±04.72	81.68±07.29	81.73±07.38
Ours	**96.30±02.87**	**96.33±02.64**	**86.56±04.67**	**86.73±04.62**

As depicted in Fig. 3, the AUC curves for both the training and testing phases across all three datasets exhibit rapid convergence, stabilizing within approximately 50 epochs. This indicates that the model efficiently learns features and quickly reaches a stable state. For the SEED and SEED-IV datasets, the test AUC values consistently exceed 0.9, demonstrating the model's excellent generalization capabilities. In contrast, for the more challenging SEED-VII dataset, the test AUC stabilizes around 0.6. The small gap between the training and testing curves across all datasets suggests that the model effectively mitigates overfitting.

Furthermore, the t-SNE visualizations in Fig. 4 highlight the quality of the domain-invariant features learned by the model. After training, the feature distributions for different emotions exhibit significantly more compact intra-class clustering and clearer inter-class separation. This demonstrates that the PR-DA model effectively aligns the joint probability distributions of the source and target domains. By creating a feature space with a more distinct structure, the model learns a powerfully discriminative representation.

TABLE II
PERFORMANCE COMPARISON OF PR-DA AND STATE-OF-THE-ART METHODS ON SEED-VII DATASET USING LEAVE-ONE-SUBJECT-OUT (LOSO) CROSS-VALIDATION.

Methods	P_{acc}	F1 score
RGNN [5]	37.49±05.44	34.52±04.83
PGCN [7]	38.67±05.41	37.28±05.49
MAET [12]	40.90±05.52	38.47±06.09
CMTCF [20]	49.13±08.22	–
Ours	**50.43±06.71**	**48.93±05.24**

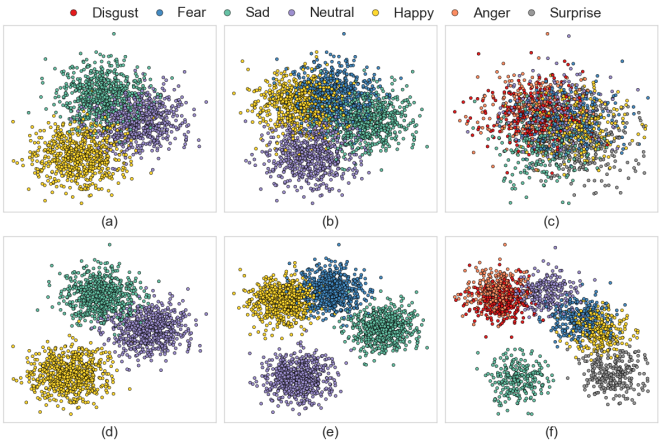

Fig. 4. t-SNE Visualization of Feature Distributions for the PR-DA Model Before (a-c) and After Training (d-f).

IV. CONCLUSION

In this study, a prototype regularization domain adaptation framework is proposed for cross-subject EEG-based emotion recognition. The connection uncertainty spatial-temporal Graph module is utilized to extract domain-invariant features by dynamically learning spatial-temporal connections. Additionally, a prototype regularization method pulls intra-class samples toward their class prototypes while separating inter-class prototypes. It then uses a bilinear interaction module to align the joint probability distributions between the source and target domains. The experimental results demonstrate that the proposed model achieves state-of-the-art performance on three EEG emotion recognition databases, SEED, SEED-IV, and SEED-VII, thereby highlighting the potential of the model.

V. ACKNOWLEDGMENT

This work was supported in part by the STI 2030-Major Projects under Grant 2022ZD0208900, in part by the National Natural Science Foundation of China under Grant 62076103, and in part by the Major Projects of Colleges and Universities in Guangdong Province under Grant 2023ZDZX2021.

REFERENCES

[1] R. Cowie, E. Douglas-Cowie, N. Tsapatsoulis, G. Votsis, S. Kollias, W. Fellenz, and J. G. Taylor, "Emotion recognition in human-computer interaction," *IEEE Signal Processing Magazine*, vol. 18, no. 1, pp. 32–80, 2001.

[2] D. Nath, Anubhav, M. Singh, D. Sethia, D. Kalra, and S. Indu, "A comparative study of subject-dependent and subject-independent strategies for EEG-based emotion recognition using lstm network," in *Proceedings of the 2020 4th International Conference on Compute and Data Analysis*, pp. 142–147, 2020.

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

[4] J.-Y. Zhu, W.-L. Zheng, and B.-L. Lu, "Cross-subject and cross-gender emotion classification from EEG," in *World Congress on Medical Physics and Biomedical Engineering, June 7-12, 2015, Toronto, Canada*, pp. 1188–1191, Springer, 2015.

[5] P. Zhong, D. Wang, and C. Miao, "EEG-based emotion recognition using regularized graph neural networks," *IEEE Transactions on Affective Computing*, vol. 13, no. 3, pp. 1290–1301, 2020.

[6] 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, 2018.

[7] M. Jin, C. Du, H. He, T. Cai, and J. Li, "PGCN: Pyramidal graph convolutional network for EEG emotion recognition," *IEEE Transactions on Multimedia*, vol. 26, pp. 9070–9082, 2024.

[8] M. Jiménez-Guarneros and G. Fuentes-Pineda, "Learning a robust unified domain adaptation framework for cross-subject EEG-based emotion recognition," *Biomedical Signal Processing and Control*, vol. 86, p. 105138, 2023.

[9] R. Zhou, Z. Zhang, H. Fu, L. Zhang, L. Li, G. Huang, F. Li, X. Yang, Y. Dong, Y.-T. Zhang, *et al.*, "PR-PL: A novel prototypical representation based pairwise learning framework for emotion recognition using EEG signals," *IEEE Transactions on Affective Computing*, vol. 15, no. 2, pp. 657–670, 2023.

[10] W. Guo, G. Xu, and Y. Wang, "Multi-source domain adaptation with spatio-temporal feature extractor for EEG emotion recognition," *Biomedical Signal Processing and Control*, vol. 84, p. 104998, 2023.

[11] W.-L. Zheng, W. Liu, Y. Lu, B.-L. Lu, and A. Cichocki, "Emotionmeter: A multimodal framework for recognizing human emotions," *IEEE Transactions on Cybernetics*, vol. 49, no. 3, pp. 1110–1122, 2018.

[12] W.-B. Jiang, X.-H. Liu, W.-L. Zheng, and B.-L. Lu, "SEED-VII: A multimodal dataset of six basic emotions with continuous labels for emotion recognition," *IEEE Transactions on Affective Computing*, vol. 16, pp. 969 – 985, 2024.

[13] D. Pan, H. Zheng, F. Xu, Y. Ouyang, Z. Jia, C. Wang, and H. Zeng, "MSFR-GCN: A multi-scale feature reconstruction graph convolutional network for EEG emotion and cognition recognition," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 31, pp. 3245–3254, 2023.

[14] H. Gao, X. Wang, Z. Chen, M. Wu, Z. Cai, L. Zhao, J. Li, and C. Liu, "Graph convolutional network with connectivity uncertainty for EEG-based emotion recognition," *IEEE Journal of Biomedical and Health Informatics*, vol. 28, pp. 5917 – 5928, 2024.

[15] J. Gu, F. Xiong, and X. Gong, "EEG-based cross subject emotion recognition based on collaborative learning and dynamic distribution adaptation," in *2024 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pp. 1960–1965, IEEE, 2024.

[16] X. Ju, J. Su, S. Dai, X. Wu, M. Li, and D. Hu, "Domain adversarial neural network with reliable pseudo-labels iteration for cross-subject EEG emotion recognition," *Knowledge-Based Systems*, vol. 316, p. 113368, 2025.

[17] Y. Wang, L. Zhang, and Y. Zhang, "Band-level adaptive fusion network for cross-subject EEG emotion recognition," *IEEE Transactions on Instrumentation and Measurement*, vol. 74, p. 2514112, 2025.

[18] K. Liu, X. Luo, W. Zhu, Z. Yu, H. Yu, B. Xiao, and W. Wu, "Enhancing EEG-based cross-subject emotion recognition via adaptive source joint domain adaptation," *IEEE Transactions on Affective Computing*, pp. 1 – 13, 2024.

[19] Y. Yang, Z. Wang, W. Tao, X. Liu, Z. Jia, B. Wang, and F. Wan, "Spectral-spatial attention alignment for multi-source domain adaptation in EEG-based emotion recognition," *IEEE Transactions on Affective Computing*, vol. 15, pp. 2012 – 2024, 2024.

[20] C. Sun, X. Wang, and L. Chen, "Channelmix-based transformer and convolutional multi-view feature fusion network for unsupervised domain adaptation in EEG emotion recognition," *Expert Systems with Applications*, vol. 280, p. 127456, 2025.