Tsinghua Science and Technology
Doi: 10.26599/TST.2026.9010060

Research Article

FSA-Net: Coarse-to-fine domain adaptation with flexible source aggregation for cross-subject EEG emotion recognition

Shuang Ran, Wei Zhong, Shuzhan Hu (✉), Long Ye, and Qin Zhang

Abstract Cross-subject emotion recognition based on electroencephalography (EEG) remains challenging due to non-stationary distribution shifts among individuals. Existing domain adaptation methods typically follow fixed paradigms, either indiscriminately merging sources, which risks negative transfer, or isolating them as independent branches, which leads to high computational redundancy. To address this dilemma, we propose a novel flexible source aggregation network (FSA-Net) for coarse-to-fine domain adaptation. Unlike static methods, our approach first introduces a similarity-based aggregation mechanism that dynamically reconfigures source domains based on their affinity to the target, effectively balancing model complexity with transfer performance. Subsequently, we construct a hierarchical alignment strategy that minimizes global discrepancy through coarse alignment using maximum mean discrepancy (MMD), as well as fine-grained subdomain adaptation incorporating contrastive learning with local MMD. This ensures that samples with identical emotion labels are clustered while distinct categories are separated in the latent space. Furthermore, a domain-aware weighted ensemble strategy is designed to integrate predictions based on the adaptation quality of each aggregated branch. Extensive experiments on the benchmark SEED and SEED-IV datasets demonstrate that FSA-Net achieves superior state-of-the-art accuracy of 90.80% and 82.74%, respectively, validating its efficacy in robust cross-subject emotion decoding.

Keywords EEG emotion recognition, cross-subject, domain adaptation, flexible source aggregation.

• Shuang Ran and Wei Zhong are with the State Key Laboratory of Media Convergence and Communication, Communication University of China, Beijing 100024, China. Shuang Ran is also with the Army Aviation Institute of PLA, Beijing, 101123, China. E-mail: rans@cuc.edu.cn, wzhong@cuc.edu.cn
• Shuzhan Hu and Long Ye are with the School of Data Science and Media Intelligence, Communication University of China, Beijing 100024, China. E-mail: hushuzhan@cuc.edu.cn, yelong@cuc.edu.cn
• Qin Zhang is with the Key Laboratory of Media Audio & Video (Communication University of China), Ministry of Education, Beijing 100024, China. E-mail: zhangqin@cuc.edu.cn
• Shuang Ran and Wei Zhong contributed equally to this work.
* To whom correspondence should be addressed.
Manuscript received: 2026-03-05; revised: 2026-04-29; accepted: 2026-06-08

1 Introduction

Emotions serve as fundamental drivers of human cognition, influencing decision-making, health, and social interactions [1–4]. As the demand for seamless human-computer interaction (HCI) grows, endowing machines with the capacity to perceive and understand human emotional states has become a paramount research objective. Among various modalities, Electroencephalography (EEG) signals are increasingly favored for emotion recognition due to their high temporal resolution and ability to objectively reflect spontaneous neural activities in the cerebral cortex [5–10], making them harder to disguise compared to facial expressions or speech.

In the nascent stages of EEG emotion recognition, research predominantly relied on traditional machine learning paradigms, coupling manual feature engineering with classifiers such as support vector machines (SVM) or random forests [11–13]. With the advent of deep learning, neural networks have demonstrated superior capability in automating feature extraction. For instance, the TSception framework was introduced [14], leveraging temporal and spatial convolutions to capture dynamic representations across time and channel domains. Similarly, EEGNet [15] was proposed as a compact convolutional neural network (CNN) architecture, utilizing depthwise and separable convolutions to efficiently filter frequency-specific information. Beyond CNNs, the potential of attention mechanisms was explored by designing the hierarchical transformer and graph convolutional networks (GCNs) to model spatial dependencies [16–19]. However, despite these architectural advancements, a critical bottleneck remains. The inherent non-stationarity and nonlinearity of EEG signals result in severe distribution shifts between individuals [20]. Consequently, models trained on source subjects often suffer from significant performance degradation when applied to unseen target subjects. Therefore, developing models with robust cross-subject generalizability is essential to bridge the gap between theoretical research and practical BCI applications [21].

To bridge this gap, domain adaptation (DA) has emerged as a promising solution, aiming to transfer knowledge by align-

Just Accepted

ing the feature distributions of labeled source domains and unlabeled target domains. However, existing DA frameworks for EEG emotion recognition face a cascade of limitations that hinder their effectiveness. At the macro level of source organization, existing approaches are often restricted to rigid paradigms that fail to capture the optimal data structure. Single-source domain adaptation (SSDA) methods typically concatenate data from all available subjects into a single super-source domain, as shown in Fig. 1(a). While computationally efficient, this one-size-fits-all approach ignores the internal distribution discrepancies among different source subjects. It often forces the alignment of conflicting distributions, leading to negative transfer. In the context of cross-subject EEG analysis, negative transfer refers to the performance degradation on the target subject caused by the forced inclusion of outlier source subjects whose neural patterns are significantly dissimilar or contradictory to the target distribution. Consequently, the inclusion of dissimilar source data degrades performance on the target [22–24]. Conversely, multi-source domain adaptation (MSDA) methods treat each source subject as a distinct domain, employing parallel branches for alignment, as shown in Fig. 1(b). Although this preserves the specific characteristics of each source, it significantly increases network complexity and overlooks the potential synergy among similar source subjects. There is a lack of a flexible strategy that balances the diversity of MSDA with the compactness of SSDA. Even when the source structure is determined, a deeper issue persists in the granularity of feature alignment. Most existing methods prioritize global distribution alignment while neglecting fine-grained, class-conditional alignment. This coarse adaptation often fails to preserve the discriminative structure of data, leading to the confusion of subdomains where samples with different labels are mixed in the latent space. Furthermore, within this alignment process, the varying transferability of different samples is frequently overlooked. Current strategies typically assume that all target domain samples and source domains are equally important for the transfer. This lack of a differentiation mechanism restricts the model from suppressing irrelevant features and focusing on the most informative components for robust emotion recognition.

To address these challenges comprehensively, this paper proposes the flexible source aggregation network (FSA-Net), a unified framework that orchestrates source data organization, feature alignment, and decision weighting to achieve robust cross-subject EEG emotion recognition. First, to optimize data organization, we depart from rigid single-source or multi-source paradigms by introducing a flexible source

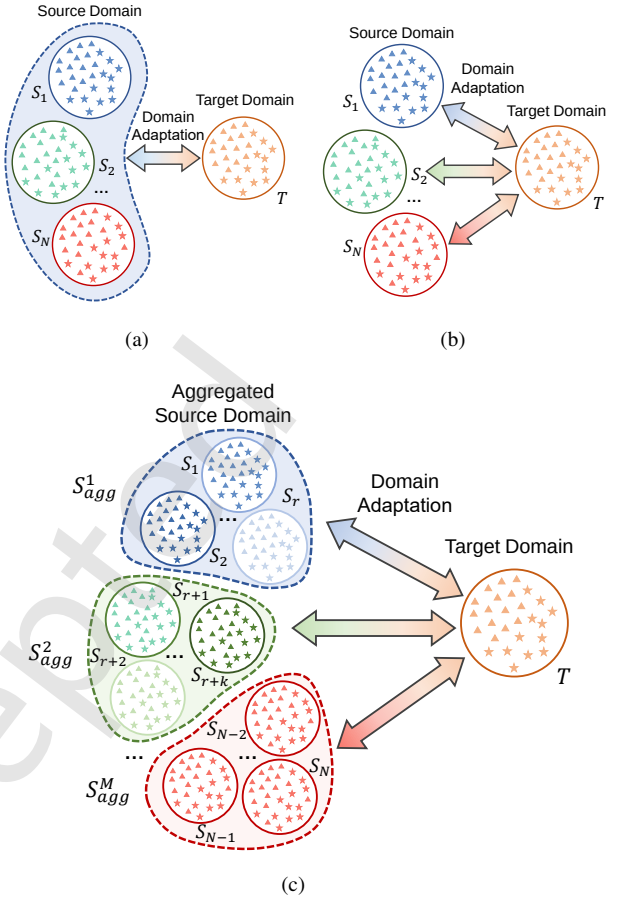


Fig. 1 Three strategies of domain adaptation: (a) single-source domain adaptation, (b) multi-source domain adaptation and (c) flexible aggregated source domain adaptation.

aggregator. This module dynamically groups source subjects based on their distributional similarity to the target domain, as shown in Fig. 1(c), effectively unifying SSDA and MSDA. This strategy balances the diversity of source data with the computational efficiency of the model. Then, to ensure precise feature adaptation, we construct a hierarchical coarse-to-fine alignment mechanism according to the previous study [25]. At the global level, maximum mean discrepancy (MMD) is employed to minimize the marginal distribution shifts between the aggregated source and target domains. At the local level, we integrate contrastive learning with local MMD (LMMD) in the reproducing kernel Hilbert space (RKHS). This dual-constraint mechanism explicitly pulls together samples of the same emotion category while pushing apart those of different categories, effectively resolving the subdomain confusion issue. Finally, to refine the decision-making process, a domain-aware weighted strategy is devised. Instead of averaging predictions, this strategy adaptively assigns higher confidence weights to classifiers trained on better-aligned aggregated domains, thereby reducing prediction bias caused

Just Accepted

by negative transfer. Extensive experiments on the SEED and SEED-IV datasets, along with comprehensive ablation studies, demonstrate the effectiveness and superiority of the proposed FSA-Net. Compared to the preliminary work that focused on validating the effectiveness of coarse-to-fine alignment using contrastive learning [25], this paper advances the framework by addressing the critical limitation of rigid source organization. The main contributions are as follows:

- We propose a unified domain adaptation framework with flexible source aggregation. Departing from single-source or multi-source paradigms, this framework dynamically reconfigures source domains based on their similarity to the target. This strategy effectively resolves the trade-off between model complexity and transfer performance, providing a generalized solution for data organization in cross-subject EEG analysis.
- We develop a hierarchical coarse-to-fine alignment mechanism incorporating contrastive learning. Beyond traditional global alignment, we integrate a subdomain-aware contrastive loss to explicitly minimize the intra-class distance while maximizing the inter-class distance in the latent space. This ensures precise distribution alignment from global statistics down to local category-specific subdomains, significantly enhancing discriminative feature learning.
- We design a domain-aware weighted ensemble strategy to ensure robust decision-making. Recognizing the varying transferability of different source domains, this mechanism adaptively integrates predictions from multiple classifiers based on their alignment quality, reducing the negative impact of outlier domains.

2 Related Works

This section reviews the evolution of domain adaptation techniques in EEG emotion recognition and explores the application of contrastive learning, highlighting the gaps that motivate our proposed FSA-Net.

2.1 Domain Adaptation for EEG Emotion Recognition

While EEG signals offer rich, objective features for emotion decoding, their inherent non-stationarity and significant inter-subject variability pose a severe challenge. Consequently, models trained on existing source data often suffer from performance degradation when generalizing to unseen target subjects due to distribution shifts. To bridge this domain gap, domain adaptation (DA) techniques [26] have been extensively employed, which can be methodologically categorized into two streams based on their alignment mechanisms [27]. The first stream comprises adversarial-based methods [28–31], primarily built upon domain adversarial neural networks (DANN) [32, 33]. These approaches employ a discriminator to distinguish domains, forcing the feature extractor to learn domain-invariant representations. The second stream focuses on statistical matching-based methods [22, 23, 34, 35], which explicitly minimize the distributional discrepancy between domains in a high-dimensional space. Due to their theoretical stability, interpretability, and ease of optimization, statistical approaches have gained widespread adoption in EEG analysis and serve as the foundation of our work. Beyond the alignment mechanism, DA frameworks are fundamentally distinguished by their source data configuration. Depending on how the source subjects are utilized, existing methods are classified into single-source domain adaptation (SSDA) and multi-source domain adaptation (MSDA).

2.1.1 Single-Source Domain Adaptation (SSDA)

SSDA strategies operate on the premise that source subjects share a coherent distribution, thereby merging all source data into a unified domain to align with the target. Early foundational works, such as the deep adaptation network (DAN) [32], utilized multi-kernel maximum mean discrepancy (MMD) [36] to minimize the global distribution shift. To further refine this alignment, subsequent research incorporated adversarial learning [28], simultaneously adapting marginal and conditional distributions to enforce latent representation similarity across subjects. Moving beyond global alignment, recent approaches have focused on class-level and structural granularity. For instance, to mitigate the misclassification of ambiguous samples, some research [37] introduced a minimum class confusion loss within a gated recurrent unit framework. Similarly, recognizing that not all samples contribute equally, some research [38] developed a sample-reweighted DA neural network (SRDANN) to prioritize high-quality data during transfer. Furthermore, the exploration of spatio-temporal dynamics has enriched SSDA. The spatio-temporal isomorphic cross-brain region interaction network (STCBI-Net) [39] was proposed to model the isomorphic consistency of brain regions, while some research [40] leveraged multi-scale hyperbolic contrastive learning (MSHCL) to capture event-related subject-invariant representations.

2.1.2 Multi-Source Domain Adaptation

In contrast to the one-size-fits-all approach of SSDA, MSDA acknowledges inter-subject heterogeneity by assigning dedicated adaptation branches to distinct source domains. A

Just Accepted

primary challenge in this paradigm is optimizing the contribution of each branch. To address this, the multi-source feature representation and alignment network (MS-FRAN) [31] was proposed, which combines a wide feature extractor with a top-h ranking mechanism to select the most relevant domain classifiers. Similarly, MEERNet [34] and MS-MDA [22] were developed, utilizing discrepancy losses to enhance the discriminative power of multi-source ensembles. Advanced MSDA frameworks also strive to disentangle domain-invariant features from domain-specific noise. Some research [35] employed MMD-based constraints to extract shared commonalities while preserving individual characteristics. By considering both domain-invariant and domain-specific features, the multi-source associate domain adaptation (MADA) network [41] was proposed for EEG emotion recognition. Furthermore, to handle outlier sources that may degrade performance, MISNet [23] with a two-stream structure was introduced to filter negative transfer, while the multi-source structural deep clustering (MSSDC) [42] was proposed to uncover the intrinsic manifold of the target domain without sacrificing class separability.

While both SSDA and MSDA have advanced the state of cross-subject emotion recognition by reducing distribution discrepancies, they represent two extremes of data organization: completely merging or completely separating sources. This motivates the need for a flexible aggregation strategy that balances sample diversity with distribution consistency.

2.2 Contrastive Learning for EEG Signal Analysis

Contrastive learning (CL) constructs discriminative representations by minimizing the distance between positive pairs while maximizing that of negative pairs. This paradigm has recently emerged as a powerful tool in brain-computer interfaces (BCI) for extracting intrinsic neural patterns [43, 44]. In broader BCI tasks such as motor imagery, CL is often utilized to handle cross-domain shifts and uncover latent geometric structures. For instance, to tackle multi-source adaptation without target data, an instance-contrastive cross-entropy loss [45] was employed to learn the intrinsic data manifold. Taking a geometric perspective, MACNet [46] was introduced, which harmonizes Euclidean and Riemannian views of EEG signals via CL to capture robust spatio-temporal features across recording sessions. In the specific context of EEG emotion recognition, CL is pivotal for suppressing individual variability and enhancing class separability. The CLISA framework [47] was proposed to explicitly align representations across different subjects, thereby minimizing inter-subject discrepancies. Furthermore, to fully leverage label information, the ECNN-C network [48] integrated supervised contrastive learning within a convolutional neural network, ensuring that EEG signals sharing the same emotion category are tightly clustered in the embedding space. Recently, some research explored a coarse-to-fine domain adaptation method utilizing contrastive learning to align subdomains [25]. However, this approach relies on fixed source domains and does not account for the optimization of source data organization, which limits its flexibility in handling diverse subjects.

Despite the advancements, critical gaps persist in current methodologies, limiting their efficacy in robust cross-subject emotion recognition. First, regarding feature alignment and sample importance, most existing frameworks prioritize global distribution matching while neglecting fine-grained, class-conditional alignment. Furthermore, they typically assign uniform importance to all target samples and source domains during the transfer process. This assumption overlooks the varying transferability of different instances, leading to negative transfer. Second, regarding source organization, existing paradigms either indiscriminately merge all source subjects into a common domain, masking internal distribution discrepancies, or isolate them as independent branches, which incurs high computational redundancy and ignores inter-subject synergies. To bridge these gaps, it is imperative to move beyond static configurations. A robust solution requires a mechanism that dynamically aggregates source domains based on distributional similarity to balance diversity and efficiency, and performs hierarchical alignment from global statistics down to local subdomains, ensuring precise knowledge transfer at different granularities.

3 Methods

In this section, we present the framework of a coarse-to-fine domain adaptation with flexible source aggregation and its data transmission process, as shown in Fig. 2. The two main challenges we address are: (1) developing a discrepancy-aware domain processing mechanism that leverages both diversity and similarity among source domains, and (2) reducing domain shift caused by non-stationary and individual differences in EEG signals. To solve these two problems, we aggregate the source domains flexibly and align the data distributions both from the source and target domains globally and those of category-relevant subdomains locally.

3.1 Overview of the Proposed Method

Let the multisource EEG data be denoted as $\mathcal{S} = \{S_1, S_2, ..., S_N\}$, where $S_i = \{(s_i^k, y_i^k)\}_{k=1}^{n_s}$ represents the

Just Accepted

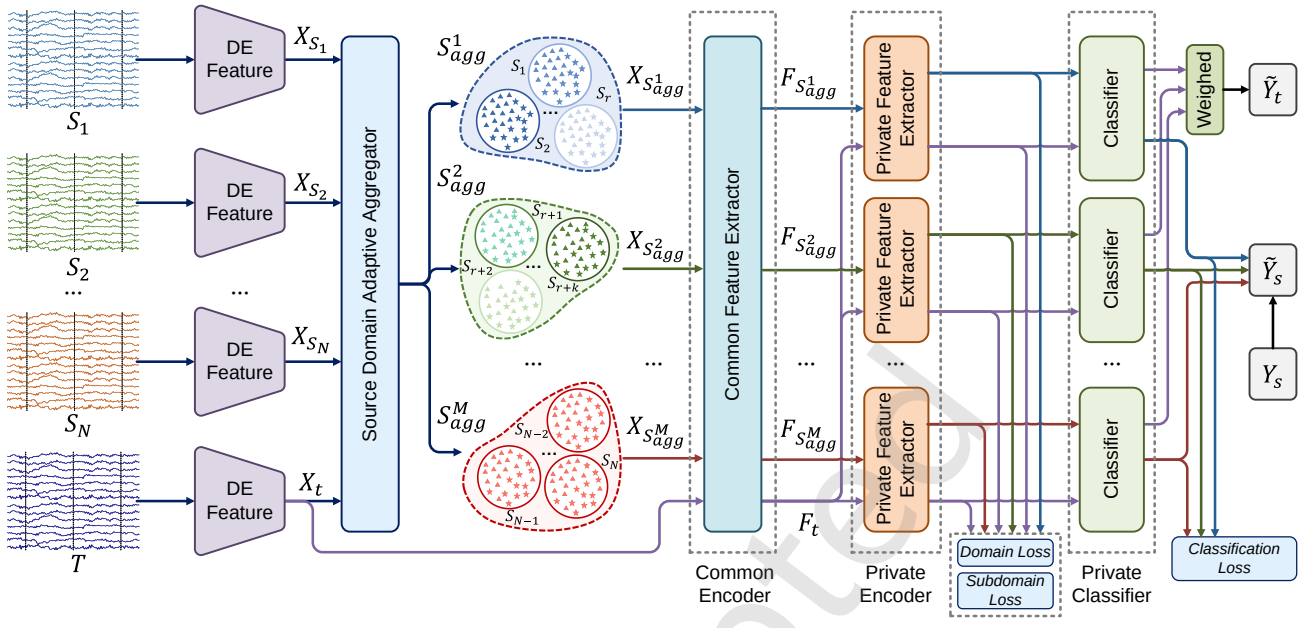


Fig. 2 The framework of the proposed coarse-to-fine domain adaptation with flexible source aggregation. Cross-subject emotion decoding was achieved through flexible source domain aggregation, hierarchical feature encoding, and weighted classifiers.

i-th source domain with n_s samples and corresponding emotion labels Y_s. The target domain is defined as $T = \{t_j\}_{j=1}^{n_t}$ containing n_t target samples. First, differential entropy (DE) features are extracted from the raw EEG signals to construct the feature sets as shown in Fig. 2. Let $X_{S_i} = \{x_{S_i}^w \in \mathbb{R}^d\}_{w=1}^{n_s'}$ and $X_T = \{x_t^z \in \mathbb{R}^d\}_{z=1}^{n_t'}$ denote the d-dimensional DE feature vectors for the i-th source domain and the target domain, respectively. The proposed pipeline proceeds as follows.

Flexible source aggregation. To optimize the transferability, we dynamically reconfigure the original N source domains based on their distributional similarity to the target. This results in M aggregated source domains, denoted as $\mathcal{S}_{agg} = \{S_{agg}^m\}_{m=1}^M$. The feature set for the m-th aggregated domain is represented as $X_{S_{agg}^m}$, containing $n'_{S_{agg}^m}$ samples.

Hierarchical feature encoding. We employ a common encoder to extract low-level, domain-invariant representations ($F_{S_{agg}^m}$ and F_t) from both the aggregated source and target inputs. Subsequently, distinct private encoders are utilized to capture high-level, domain-specific features for each aggregated branch, preserving the unique characteristics of the grouped data.

Optimization and inference. The network is trained by minimizing a joint loss function consisting of the following parts. A classification loss ($\mathcal{L}_{class}$) using source labels Y_s and predicted labels $\tilde{Y}_s$ to ensure distinct decision boundaries. A global alignment loss ($\mathcal{L}_{domain}$) to minimize the marginal distribution shift between $X_{S_{agg}^m}$ and T. A local contrastive loss ($\mathcal{L}_{sub}$) to align category-relevant subdomains, ensuring intra-class compactness.

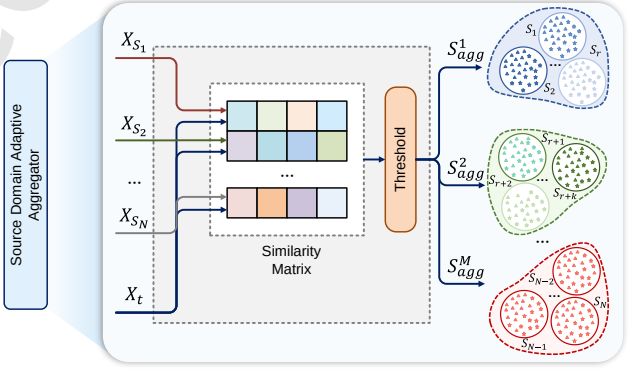


Fig. 3 The details of the source domain adaptive aggregator.

Finally, during the inference phase, the predictions from the private classifiers are integrated using a domain-aware weighted strategy, which assigns higher importance to branches with superior adaptation quality.

3.2 Source Domain Adaptive Aggregator

For the problems of ignoring the differences within source domains in SSDA and the high computational complexity in MSDA, we design the source domain adaptive aggregator to aggregate the source domains adaptively, as shown in Fig. 3. It has been shown that source domains exhibiting higher feature similarity to the target domain offer more reliable

Just Accepted

information for positive transfer. Consequently, we quantify the transferability of the i-th source domain S_i by calculating the similarity vector A^i between the i-th source domain X_{S_i} and the target domain X_t. Specifically, for each target sample x_t^z $(z = 1, 2, ..., n_t')$, its average cosine similarity with all samples in X_{S_i} is computed as:

$$A^i = \text{Similarity}(X_{S_i},\ X_t) = \frac{1}{n_s'} \sum_{w=1}^{n_s'} \left|\cos(x_{S_i}^w, x_t^z)\right|. \quad (1)$$

The choice of cosine similarity over other density-based distribution measures is motivated by its computational efficiency and its ability to robustly capture the directional alignment of high-dimensional feature vectors, which is highly effective for an initial quantification of subject affinity before formal domain alignment. Consequently, $A^i = [A_1^i, A_2^i, ..., A_{n_t'}^i] \in \mathbb{R}^{n_t'}$ forms a similarity distribution vector for the i-th source domain. To obtain a global transferability score for sorting, we compute the norm of this vector, denoted as $a^i = ||A^i||$. This yields a global affinity set $\mathbf{A} = \{a^1, a^2, ..., a^N\}$ for all N source domains. To construct the aggregated domains, we perform a ranking-based segmentation. First, the affinity set $\mathbf{A}$ is sorted in descending order to form $\mathbf{A}_{sort}$. We then iterate through this sorted sequence to evaluate the transferability gap between adjacent source subjects. Guided by a smoothness assumption, if the difference between two adjacent scores is within a pre-defined threshold th (i.e., $a_{sort}^j - a_{sort}^{j+1} \leqslant th$), the corresponding subjects are deemed to share homogeneous transfer characteristics and are merged into the same aggregated domain S_{agg}^m. Conversely, if the gap exceeds th, a significant distribution shift boundary is detected, triggering the initialization of a new aggregated domain S_{agg}^{m+1}.

This recursive sequential clustering automatically determines the optimal number of aggregated domains, M. Source subjects at the tail of the sorted list that fail to group with others (isolated outliers with massive domain shifts) are discarded to prevent negative transfer. Crucially, the threshold th acts as a granularity controller. A sufficiently large th merges all diverse sources into a single domain (M approaches 1), recovering the SSDA paradigm, while a near-zero th preserves individual distinctions (M approaches N), defaulting to the MSDA paradigm.

3.3 Common Encoder

Despite the significant non-stationarity and individual variability in EEG signals, underlying neurophysiological patterns regarding emotional processing remain shared across human brains. Capturing these universal patterns is crucial for establishing a baseline for subsequent adaptation. Therefore, we design a common encoder E_c to project the DE features from both the aggregated source and target domains into a shared latent space. E_c is composed of three stacked linear layers with LeakyReLU activation functions. This architecture is designed to extract low-level, generic representations that filter out subject-specific noise while retaining emotional semantics. Formally, let $X_{S_{agg}^m}$ and X_t denote the input feature sets. The common feature extraction process is defined as:

$$F_{S_{agg}^m} = E_c(X_{S_{agg}^m}),\ m = 1, 2, ..., M, \quad (2)$$

$$F_t = E_c(X_t), \quad (3)$$

where $F_{S_{agg}^m}$ and F_t represent the extracted low-level common features. These shared representations serve as the foundational input for the subsequent domain-specific adaptation in the private encoders.

3.4 Private Encoder

Following the common encoder, we deploy a bank of private encoders to capture high-level, discriminative semantic representations. The common encoder E_c is intended to extract universal low-level features that are fundamentally shared across individuals. However, such generic features often lack the fine-grained resolution needed to align specific subdomains. The private encoders E_{pri} are thus expected to refine these common features into specialized, high-level semantic representations for each aggregated branch, facilitating a more precise mapping. To accommodate the diversity of the M aggregated source domains, we instantiate M parallel branches. Each branch employs a linear layer with LeakyReLU activation to map the low-level common features into a domain-specific latent space. Simultaneously, a target private encoder is employed to extract target-specific features. The high-level feature extraction is formulated as:

$$F'_{S_{agg}^m} = E_{pri}^m(F_{S_{agg}^m}),\ m = 1, 2, ..., M, \quad (4)$$

$$F'_t = E_{pri}(F_t). \quad (5)$$

3.4.1 Global alignment

Merely extracting features is insufficient for robust transfer, and the marginal distributions between domains must be aligned in the latent space. To achieve this, we introduce MMD as a global metric. MMD measures the discrepancy between two probability distributions by mapping them into a reproducing kernel Hilbert space ($\mathcal{H}$) [36].

Just Accepted

$$\begin{aligned}&\mathrm{MMD}^2(p^m, q)\\&= \left\| \mathbb{E}_{f'_{S^m_{agg}} \sim p^m} \left[\phi\left(f'_{S^m_{agg}} \right) \right] - \mathbb{E}_{f'_t \sim q} \left[\phi\left(f'_t \right) \right] \right\|^2_{\mathcal{H}},\end{aligned} \tag{6}$$

where the distribution of m-th aggregated source domain $F'_{S^m_{agg}}$ is p^m and that of the target domain F'_t is q, and $\phi(\cdot) : \mathcal{F} \to \mathcal{H}$ denotes the feature map function associated with the kernel. Given the significant variations in EEG signals, we utilize MMD to explicitly enforce the global alignment of marginal distributions between each aggregated source branch and the target domain. Formally, we define the global domain adaptation loss $\mathcal{L}_{domain}$ by minimizing the MMD distance between the high-level features of each aggregated source domain and the target domain:

$$\begin{aligned}\mathcal{L}_{domain} &= \sum_{m=1}^{M} \mathrm{MMD}^2(F'_{S^m_{agg}}, F'_t)\\&= \sum_{m=1}^{M} \left\| \frac{1}{n'_{S^m_{agg}}} \sum_{i=1}^{n'_{S^m_{agg}}} \phi(f'^i_{S^m_{agg}}) - \frac{1}{n'_t} \sum_{j=1}^{n'_t} \phi(f'^j_t) \right\|^2_{\mathcal{H}},\end{aligned} \tag{7}$$

While minimizing $\mathcal{L}_{domain}$ successfully reduces the marginal distribution shift, it operates on the domain as a whole, ignoring the fine-grained structure of emotion categories. This coarse alignment often leads to subdomain confusion, where samples of different emotions from the target domain may be mapped incorrectly near the decision boundaries of the source domain. To bridge this gap, local MMD provides statistical alignment for category-specific centroids, while contrastive learning enforces instance-level discriminability. This hierarchical strategy ensures that the feature space is not only globally aligned but also locally well-separated.

3.4.2 Local alignment

While global alignment minimizes marginal distribution shifts, it overlooks the conditional distributions of specific emotion categories (subdomains). This often results in negative transfer, where samples with different labels are mapped close together. To address this, we perform fine-grained alignment using local MMD (LMMD) [49] on the high-level latent features. Since target labels are unavailable, we utilize the predictions from the private classifiers as pseudo-labels. Let $F'^c_{S^m_{agg}}$ and F'^c_t denote the set of high-level feature vectors (output of private encoders) belonging to emotion category $c \in \{1, 2, ..., Cl\}$ in the m-th aggregated source and target domains, respectively. The LMMD metric measures the discrepancy between these class-conditional distributions in the RKHS:

$$\begin{aligned}&\mathrm{LMMD}^2(p^m_c, q^c) =\\&\mathbb{E}_c \left\| \mathbb{E}_{f'^c_{S^m_{agg}} \sim p^m_c} \left[\phi\left(f'^c_{S^m_{agg}} \right) \right] - \mathbb{E}_{f'^c_t \sim q^c} \left[\phi\left(f'^c_t \right) \right] \right\|^2_{\mathcal{H}},\end{aligned} \tag{8}$$

where the distribution of c-th subdomain of the $F'^c_{S^m_{agg}}$ is p^m_c and that of the F'^c_t is q^c. This ensures that the centroid of each emotion cluster in the source domain is aligned with its counterpart in the target domain.

To further enhance the discriminative power, we integrate LMMD into a supervised contrastive learning framework. While LMMD aligns distribution statistics, contrastive learning explicitly enforces intra-class compactness and inter-class separability at the instance level. We construct a unified feature set $\mathcal{U}$ containing features from both the aggregated source domains (with ground-truth labels) and the target domain (with pseudo-labels). We define a subdomain contrastive loss $\mathcal{L}_{sub}$:

$$\mathcal{L}_{sub} = \sum_{i \in \mathcal{U}} \frac{-1}{|P(i)|} \sum_{p \in P(i)} \log \frac{\exp(\frac{k(f'_i, f'_p)}{\tau})}{\sum_{a \in A(i)} \exp(\frac{k(f'_i, f'_a)}{\tau})}, \tag{9}$$

where f'_i represents the high-level feature of an anchor sample, and $A(i) \equiv \mathcal{U} \backslash \{i\}$ is the set of all other samples. $P(i)$ denotes the set of positive samples sharing the same emotion category (or pseudo-label) as i. Specifically, we employ a kernel-based similarity metric $k(\cdot)$ consistent with the LMMD formulation. $\tau = 0.07$ is the temperature parameter. This value is selected based on common practice in supervised contrastive learning literature, as it provides an optimal scaling for the logits to balance the sharpness of the distribution and the stability of gradient updates. By minimizing $\mathcal{L}_{sub}$, the model explicitly shapes the feature space to form tight, well-separated emotion clusters, preventing the mixing of subdomains.

3.5 Weighted Classifiers

Following the private encoders, we instantiate M parallel private classifiers corresponding to the aggregated source domains. Each classifier, composed of a fully connected layer and a Softmax function, maps the domain-specific features $F'_{S^m_{agg}}$ to probability distributions over the emotion categories. To ensure discriminative capability, we minimize the classification loss $\mathcal{L}_{cl}$ using cross-entropy across all labeled samples in the aggregated source domains:

$$\mathcal{L}_{cl} = \sum_{m=1}^{M} \mathbb{E}_{(s,y) \sim S^m_{agg}} \left[\mathcal{J}(\hat{y}^m, y^m) \right], \tag{10}$$

where $\hat{y}^m$ is the prediction of the m-th classifier and $\mathcal{J}(\cdot)$ denotes the standard cross-entropy loss function. Combining the classification task with the proposed coarse-to-fine

Just Accepted

alignment mechanisms, the overall optimization objective is formulated as:

$$\mathcal{L} = \mathcal{L}_{cl} + \lambda_1 \mathcal{L}_{domain} + \lambda_2 \mathcal{L}_{sub}, \tag{11}$$

where λ_1 and λ_2 are trade-off hyperparameters governing the contribution of global and local alignment, respectively. To stabilize the training process, we adopt a dynamic scheduling strategy to gradually increase the weight of adaptation losses:

$$\lambda_1 = \lambda_2 = \frac{2}{1 + \exp(-\alpha \times i_{iter}/n_{iter})} - 1, \tag{12}$$

where $\alpha = 10$ is a scaling factor. This strategy suppresses noisy gradients from the alignment modules during the early stages, allowing the network to focus on learning discriminative source features before enforcing strict domain invariance.

Since the aggregated source domains exhibit varying degrees of distributional discrepancy relative to the target, treating their corresponding classifiers $\{C_m\}_{m=1}^{M}$ equally is suboptimal. Outlier domains with large shifts may dominate the decision, leading to negative transfer. To mitigate this, we propose a domain-aware weighted strategy that adaptively assigns higher confidence to classifiers trained on better-aligned domains. We quantify the reliability of the m-th classifier using the inverse of the global distribution distance between the m-th aggregated source domain and the target domain. Utilizing the global MMD metric, the unnormalized importance score D^m_{MMD} is formulated as:

$$D^m_{\text{MMD}} = \frac{1}{\text{MMD}^2(F'_{S^m_{agg}}, F'_t) + \epsilon}, \tag{13}$$

where ϵ is a small constant to ensure numerical stability. Subsequently, the normalized weight w_m for the m-th classifier is obtained by:

$$w_m = \frac{D^m_{\text{MMD}}}{\sum_{k=1}^{M} D^m_{\text{MMD}}}, \ k = 1, 2, ..., M. \tag{14}$$

This mechanism ensures that source domains closer to the target in the RKHS contribute more significantly to the final decision.

During the testing phase, the final label for a target sample is predicted by the weighted consensus of all private classifiers:

$$\hat{Y}_t = \sum_{m=1}^{M} w_m C_m(F'_t), \tag{15}$$

where C_m is the m-th classifier trained by the corresponding aggregated source domain.

In summary, the proposed FSA-Net workflow operates as follows: First, the network accepts EEG data from N source subjects and one target subject. A similarity-based aggregation mechanism dynamically orchestrates the original source data into M optimized aggregated domains. Subsequently, a common encoder extracts universal low-level features across

Algorithm 1 Training and Inference of FSA-Net

Require: Source data $X_S = \{X_{S_i}\}_{i=1}^{N}$ and labels $Y_S = \{Y_{S_i}\}_{i=1}^{N}$; Target data X_t; Max iterations N_{iter}; Learning rate η; Threshold th;

Ensure: Predicted target labels $\hat{Y}_t$;

// Phase 1: Flexible Source Aggregation
Calculate similarity vector A^i for each source domain
Compute affinity scores: $a^i = \|A^i\|$, for $i = 1, ..., N$
Sort affinity scores in descending order: $\mathbf{A}_{sort}, \mathbf{idx} = \text{sort_descending}(\{a^i\})$
Initialize $m = 1$, $X_{S^1_{agg}} = X_{S_{\mathbf{idx}[1]}}$
for $k = 2$ to N **do**
  Let current source index $i = \mathbf{idx}[k]$
  if $(\mathbf{A}_{sort}[k-1] - \mathbf{A}_{sort}[k]) \leqslant th$ **then**
    Merge source: $X_{S^m_{agg}} \leftarrow X_{S^m_{agg}} \cup X_{S_i}$
  else
    Create new group: $m = m + 1$
    $X_{S^m_{agg}} = X_{S_i}$
  end if
end for
Obtain M aggregated source domains $\{S^m_{agg}\}_{m=1}^{M}$
// Phase 2: Training
for $iter = 1$ to N_{iter} **do**
  Sample mini-batch B_t from X_t
  Initialize total loss $\mathcal{L}_{total} = 0$
  for $m = 1$ to M **do**
    Sample mini-batch B^m_s from $X_{S^m_{agg}}$
    Forward Pass:
    $F_S = E_c(B^m_s), \quad F_t = E_c(B_t)$
    $F'_S = E^m_{pri}(F_S), \quad F'_t = E_{pri}(F_t)$
    Predictions: $\hat{Y}^m_s = C_m(F'_S)$
    Compute Losses:
    $\mathcal{L}^m_{cl} \rightarrow$ Eq. (10), $\mathcal{L}^m_{domain} \rightarrow$ Eq. (7), $\mathcal{L}^m_{sub} \rightarrow$ Eq. (9)
    Update dynamic weights λ_1, λ_2
    $\mathcal{L}_{total} = \mathcal{L}_{total} + (\mathcal{L}^m_{cl} + \lambda_1 \mathcal{L}^m_{domain} + \lambda_2 \mathcal{L}^m_{sub})$
  end for
  Backward Pass & Update:
  Update parameters $\Theta = \{E_c, E_{pri}, C\}$:
  $\Theta \leftarrow \Theta - \eta \nabla_\Theta \mathcal{L}_{total}$
end for
// Phase 3: Testing & Inference
Calculate classifier weights w_m via Eq. (13) and (14) using final MMD values
Extract target features: $F'_t = E_{pri}(E_c(X_t))$
Ensemble Prediction: $\hat{Y}_t = \sum_{m=1}^{M} w_m C_m(F'_t)$ (Eq. 15)
return $\hat{Y}_t$

Just Accepted

all domains. These features are then fed into parallel private encoders to capture domain-specific representations, which are optimized via a joint objective function comprising classification, global MMD alignment, and local contrastive subdomain alignment. Finally, a domain-aware weighted ensemble strategy integrates the predictions from all private classifiers using the pre-determined weights to generate the final decision for the target domain. The detailed procedural workflow is presented in Algorithm 1.

4 Experiments

To validate the effectiveness of FSA-Net, we conducted extensive evaluations on two widely-used benchmarks: SEED and SEED-IV. This section details the experimental setup, including dataset protocols, preprocessing steps, and implementation specifics.

4.1 Datasets and Settings

4.1.1 SEED

The SEED dataset [50] is a standard benchmark for three-category emotion recognition (positive, neutral, negative). It contains EEG recordings from 15 subjects (7 males, 8 females) collected using a 62-channel ESI NeuroScan system at a 1,000 Hz sampling rate. The stimuli consist of 15 film clips, each lasting approximately 4 minutes. The data collection was conducted across three separate sessions for each subject, ensuring temporal variability. In our experiments, the pre-extracted DE features are utilized. For each session, the EEG signals are segmented into 3,394 samples per subject.

4.1.2 SEED-IV

To evaluate performance on fine-grained emotions, we utilize the SEED-IV dataset [51], which shares the same acquisition hardware as SEED but introduces higher complexity. It includes four emotion categories: happy, neutral, sad, and fear. The stimuli comprise 24 movie clips (6 per category), each approximately 2 minutes long. Similar to SEED, 15 subjects participated in three sessions held on different days. Using the same DE feature extraction protocol, this dataset provides 822 samples per subject per session.

4.1.3 Implementation Details

The proposed FSA-Net is implemented using PyTorch. The common encoder is designed as a 3-layer multilayer perceptron to map the 310-dimensional input to a 64-dimensional common feature space. The layer sizes are set as 256, 128, and 64, respectively. This 3-layer structure provides a balanced depth to transform the 310-dimensional DE features into a latent space without the risk of severe overfitting or excessive computational cost. Subsequently, the private encoder maps these features to a 32-dimensional latent space. Finally, the private classifier consists of a single fully connected layer projecting the 32-dimensional features to the label space. LeakyReLU (negative slope = 0.01) is employed as the activation function for all hidden layers.

Prior to training, we apply electrode-wise normalization to standardize the source and target data. The network is trained using the Adam optimizer with a learning rate of 0.01 and a batch size of 16. The training process spans 100 epochs. All experiments are conducted on an NVIDIA RTX 1080Ti GPU.

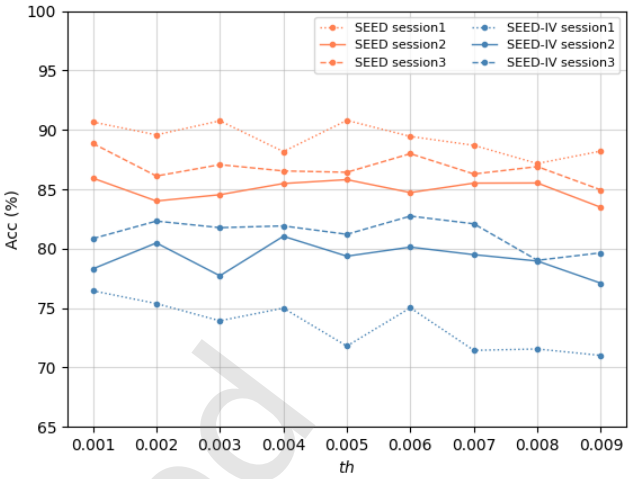


Fig. 4 Evaluations of different settings of th in terms of accuracy on the SEED and SEED-IV datasets.

4.2 Hyper-Parameter Evaluation

The threshold parameter th plays an important role in the FSA-Net framework by governing the granularity of source aggregation. As formulated in Section 3.2, a smaller th preserves the independence of source domains, while a larger th promotes aggressive merging. To identify the optimal operating point, we performed a sensitivity analysis on th to explore the impact of varying settings on the accuracy of the SEED and SEED-IV datasets. It can be seen from Fig. 4 that, when the value of th is small, it is closer to multi-source domain adaptation, and the model performance is superior to the situation when the value of th is large and aligns more closely with single-source domain adaptation. The intuition behind selecting th lies in balancing model complexity with subject diversity. A small th treats sources as independent to preserve heterogeneity, while a large th merges similar sources to mitigate computational redundancy and noise. Based on Fig. 4, setting $th = 0.005$ and $th = 0.006$ have respectively achieved peak recognition performance on the

Just Accepted

Table 1 The comparisons of the proposed method and baseline models in three sessions on the SEED dataset.

Sub.	DDC [52]			MS-MDA [22]			Ours		
	sess1	sess2	sess3	sess1	sess2	sess3	sess1	sess2	sess3
djc	83.68	75.40	77.87	84.97	80.94	76.25	91.07	79.46	80.05
jl	86.98	78.31	92.01	89.16	92.43	96.67	98.58	89.75	97.55
jj	75.81	76.37	80.38	95.64	73.33	61.79	93.05	73.10	89.83
lqj	80.70	69.33	81.79	84.44	84.35	69.77	87.18	87.86	82.35
ly	83.91	62.31	79.17	97.20	72.54	80.20	93.02	68.27	80.35
mhw	85.06	93.49	98.99	82.00	97.44	99.94	99.94	99.26	99.94
phl	76.63	82.09	82.26	93.49	79.55	76.07	86.74	79.76	76.72
sxy	84.27	86.71	75.00	77.11	94.58	78.64	83.71	98.70	74.69
wk	79.67	84.00	89.07	76.40	87.57	94.93	87.51	81.97	86.92
ww	94.25	77.37	78.99	88.07	84.59	80.35	99.94	85.86	84.83
wsf	77.40	85.71	74.22	84.56	85.36	74.87	90.07	94.49	83.21
wyw	86.71	68.35	84.15	69.39	78.93	93.43	71.24	72.01	91.10
xyl	78.52	69.06	76.66	86.65	73.19	85.12	98.32	89.80	87.63
ys	82.67	89.83	92.84	73.89	79.14	96.43	81.88	88.89	99.94
zjy	80.52	92.72	88.80	83.74	96.20	97.88	99.73	97.97	81.35
Avg.	82.45	79.40	83.48	84.45	84.01	84.16	**90.80**	**85.81**	**86.43**
Std.	**4.66**	9.16	**7.07**	7.63	**8.03**	11.34	7.86	9.54	7.70

Avg.: average accuracy, Std.: standard deviation, Sub.: subject name. The accuracy in the table is presented as a percentage (%).

Table 2 The comparisons of the proposed method and baseline models in three sessions on the SEED-IV dataset.

Sub.	DDC [52]			MS-MDA [22]			Ours		
	sess1	sess2	sess3	sess1	sess2	sess3	sess1	sess2	sess3
cz	75.20	72.60	72.14	84.14	68.51	60.95	88.25	79.21	80.29
ha	71.33	62.62	64.72	71.21	61.06	63.26	79.32	60.82	60.16
hql	56.99	74.88	65.81	75.44	74.52	74.09	69.45	78.25	78.83
ldy	60.28	67.91	66.18	61.46	67.55	69.10	83.90	79.45	83.70
ly	64.16	64.18	73.24	66.75	74.04	73.11	78.26	63.58	91.12
mhw	73.91	77.04	63.14	71.21	79.93	83.82	69.68	78.61	79.68
mz	58.05	74.88	77.86	67.69	82.33	59.61	64.98	86.78	86.25
qyt	61.22	87.02	76.64	82.96	100.00	90.27	77.20	94.47	78.71
rx	68.98	81.37	74.94	69.92	92.91	82.36	81.79	85.58	86.25
tyc	66.39	60.82	77.01	67.92	61.18	74.94	76.38	84.85	82.73
whh	55.46	71.27	84.79	81.43	75.72	67.40	71.33	85.82	96.59
wll	49.82	68.27	57.91	42.07	59.98	55.23	63.69	65.14	81.87
wq	72.03	80.77	82.48	81.79	91.47	86.25	78.26	89.06	85.16
zjd	61.81	70.67	78.83	69.45	83.05	72.02	62.98	91.23	88.69
zjy	71.21	73.80	72.26	67.20	76.32	63.63	79.79	79.09	81.02
Avg.	64.46	72.54	72.53	70.71	76.57	71.74	**75.02**	**80.13**	**82.74**
Std.	**7.33**	**7.02**	**7.34**	10.15	11.65	10.10	7.37	9.72	7.70

Avg.: average accuracy, Std.: standard deviation, Sub.: subject name. The accuracy in the table is presented as a percentage (%).

first session of SEED and the third session of SEED-IV datasets. Therefore, we use these settings to evaluate the effectiveness of the network in the following experiments.

4.3 Comparisons with Baseline

To validate the effectiveness of the proposed flexible source aggregation strategy, we compare FSA-Net against two representative baseline models: DDC [52] and MS-MDA [22]. These models are selected because they share a similar feature extractor and classifier architecture as our method but employ fixed source organizations. DDC represents the standard SSDA paradigm, while MS-MDA represents the MSDA paradigm. This setup ensures that any performance gain is attributed to our aggregation and alignment mechanism rather than differences in the backbone network. The quantitative results on the SEED and SEED-IV datasets, conducted under a leave-one-subject-out protocol, are summarized in Table 1 and Table 2, respectively.

As evident from the tables, FSA-Net consistently outperforms both baselines across all three sessions on both datasets. On the SEED dataset, our method achieves an average accuracy of 90.80% in Session 1, surpassing the single-source baseline (DDC) by 8.35% and the multi-source baseline (MS-MDA) by 6.35%. On the SEED-IV dataset, the advantage is more pronounced on this challenging four-class task. In Session 3, FSA-Net reaches 82.74%, yielding improvements of 10.21% and 11.00% over DDC and MS-MDA, respectively. In addition, it can be seen from the Tables that for the same subject and session, in most cases, the performance of our method is usually better than the baseline models. These results empirically validate that our method effectively bridges the gap between SSDA and MSDA. By flexibly aggregating source domains, FSA-Net mitigates the negative transfer caused by conflicting distributions in DDC and reduces the redundancy inherent in MS-MDA. Regarding stability, although the standard deviation of our method is marginally higher in certain cases, this is an acceptable trade-off given the substantial boost in mean accuracy. This variation reflects the inherent difficulty in adapting to subjects with extreme distribution shifts, yet the overall high accuracy confirms the reliability of FSA-Net for cross-subject emotion decoding. Notably, FSA-Net exhibits a more pronounced performance margin on SEED-IV than on SEED. This observation suggests that as the complexity of the recognition task increases and the decision boundaries become more intricate, the benefits of our hierarchical coarse-to-fine alignment and flexible source filtering become more critical. By explicitly refining subdomains, FSA-Net effectively manages the overlap of ambiguous emotional states in complex multi-class scenarios. Based on these findings, the subsequent comparison results (Section 4.4), ablation studies (Section 4.5), and visualizations (Section 4.6) will be conducted using the best-performing sessions to provide a deeper analysis.

4.4 Comparisons with Competing Methods

To comprehensively assess the performance of FSA-Net, we compare it against a wide range of state-of-the-art (SOTA) algorithms on the SEED and SEED-IV datasets. As summarized

Just Accepted

Table 3 Performance comparisons with competing methods on the SEED and SEED-IV datasets.

Domain	Methods	SEED		SEED-IV	
		Avg. (%)	Std. (%)	Avg. (%)	Std. (%)
Single-source	DCORAL [53]	66.39	7.55	51.85	7.30
	DDC [52]	83.48	7.07	72.54	**7.02**
	DAN [54]	83.81	8.56	58.87	8.13
	BIHDM [55]	87.07	7.14	69.03	8.66
	JDA [28]	88.28	11.44	70.83	10.25
	CLISA [47]	86.40	6.40	-	-
	DA-CapsNet [56]	84.63	9.09	-	-
	STCBI-Nets [39]	90.21	6.32	75.17	8.82
	MSHCL [40]	89.30	4.40	72.25	9.37
	LGDAAN [57]	89.09	7.76	74.12	8.45
Multi-source	MEER-NET [34]	87.10	**2.00**	71.00	12.10
	MS-MDA [22]	84.45	7.63	67.96	11.94
	MSFRAN [31]	85.04	6.02	-	-
	MWACN [58]	89.30	9.18	74.60	10.77
	MFA-LR [59]	89.11	7.72	74.99	12.10
	MADA [41]	86.16	7.87	59.29	13.65
	MISNet [23]	88.80	6.20	74.60	9.30
	MSSDC [42]	88.02	9.69	71.49	13.58
FSA	Ours	**90.80**	7.86	**82.74**	7.70

in Table 3, the competing methods are categorized into three groups: SSDA, MSDA, and our proposed flexible aggregated source DA. The results for DCORAL [53], DDC [52], and MS-MDA [22] are reproduced using their official open-source implementations, while results for other baselines are cited directly from their respective original publications.

From Table 3, the performance of the proposed FSA-Net achieves an average accuracy of 90.80% on the SEED dataset, outperforming the best-performing SSDA and MSDA methods. On the SEED-IV dataset, the performance of FSA-Net reaches 82.74%, demonstrating an improvement over existing approaches. The improvements of FSA-Net stem from two synergistic advantages. Unlike SSDA which risks negative transfer, or MSDA which suffers from redundancy, our flexible aggregation strategy optimally balances diversity and consistency. This allows the model to selectively leverage the most beneficial source subjects. In addition, the integration of multi-metric adaptation ensures that the model not only aligns domain distributions but also preserves the discriminative structure of emotion subdomains. Furthermore, regarding stability, FSA-Net exhibits competitive standard deviations, particularly on the complex SEED-IV dataset. This indicates that our discrepancy-aware mechanism effectively overcomes individual distribution shifts, offering a robust solution for practical BCI applications.

Table 4 Ablation study for subject-independent recognition accuracy on the SEED and SEED-IV datasets.

Model	SEED	SEED-IV
all components	$\mathbf{90.80 \pm 7.86}$	$\mathbf{82.74 \pm 7.70}$
w/o* $\mathcal{L}_{domain} + W_m$	87.16 ± 8.19	77.86 ± 10.03
w/o $\mathcal{L}_{sub} + W_m$	88.01 ± 9.07	79.84 ± 10.65
w/o $\mathcal{L}_{sub}$	88.35 ± 8.09	81.21 ± 7.79
w/o W_m	88.79 ± 10.55	75.36 ± 11.40
only $\mathcal{L}_{cl}$	72.77 ± 7.53	65.63 ± 9.26

* w/o denotes the ablated model trained without the following components.

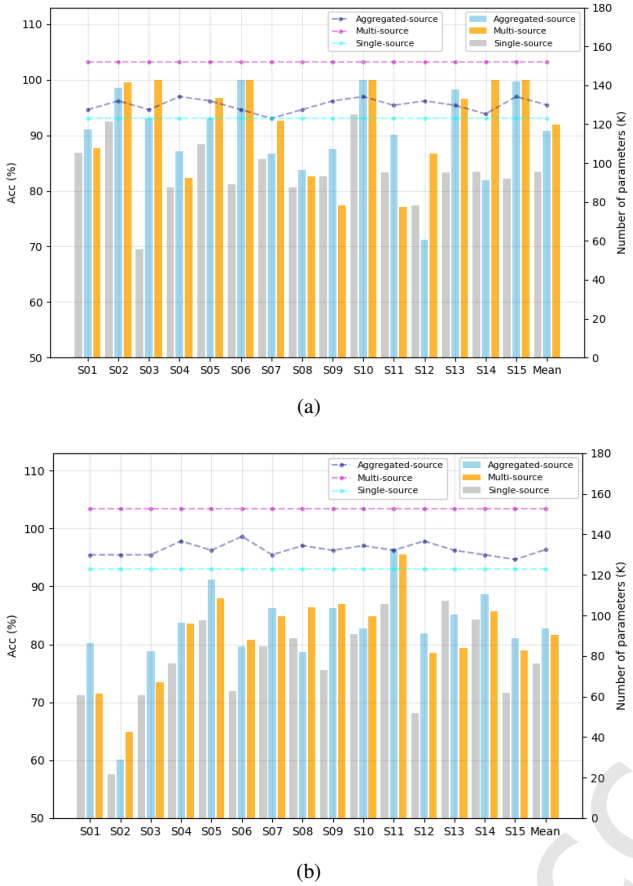
4.5 Ablation Study

To empirically quantify the contribution of each constituent module within FSA-Net, we conducted a component-wise ablation study on the SEED and SEED-IV datasets. The results are detailed in Table 4. Note that the domain-aware weighting strategy (W_m) is intrinsically derived from the global MMD metric. Therefore, the configuration involving W_m without the corresponding global alignment ($\mathcal{L}_{domain}$) is invalid and excluded. The complete FSA-Net architecture achieves the highest accuracies of $90.80 \pm 7.86\%$ and $82.74 \pm 7.70\%$, significantly outperforming all ablated variants. In contrast, the baseline model trained solely with classification loss ($\mathcal{L}_{cl}$) suffers the most severe degradation. This fundamental gap emphasizes the necessity of domain adaptation in mitigating the massive distribution shifts inherent in EEG signals. Analyzing the alignment granularity, we observe that removing the global alignment component ($\mathcal{L}_{domain} + W_m$) results in a sharper performance drop compared to removing the local component ($\mathcal{L}_{sub}$). This suggests that while fine-grained subdomain alignment refines the decision boundaries, minimizing the global marginal distribution shift remains the primary driver for successful transfer. Crucially, the contribution of the weighting strategy (W_m) is more pronounced on the SEED-IV dataset than on SEED. As indicated by the higher standard deviation, SEED-IV exhibits greater inter-subject variability and complexity. In such challenging scenarios, the negative transfer from outlier source domains is more detrimental. Therefore, the ability of W_m to adaptively suppress low-quality sources and prioritize beneficial ones becomes useful in stabilizing the model's performance.

4.6 Additional Evaluations

In this section, we conduct a comprehensive assessment of FSA-Net from three perspectives: the trade-off between model complexity and performance under different source configurations, the visualization of latent feature distributions, and the discriminative ability across specific emotion categories.

Just Accepted



(a)

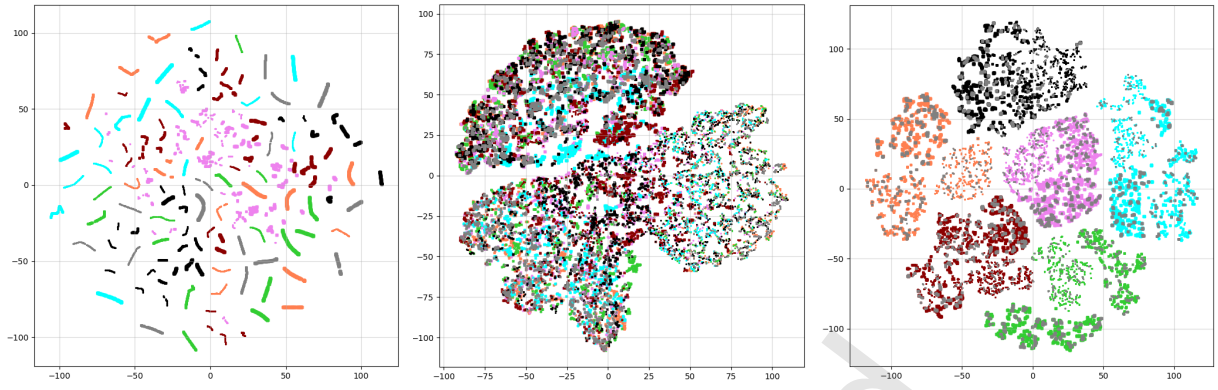
(b)

Fig. 5 The evaluations on the accuracy and number of parameters for three types of domain adaptation models on the datasets of (a) SEED and (b) SEED-IV.

(1) Balancing performance and complexity. As theoretically established in Section 3, SSDA and MSDA serve as special boundary cases of our flexible aggregated source DA, controlled by the threshold th. To empirically validate the superiority of our dynamic aggregation strategy, we compare these three paradigms across individual subjects on the SEED and SEED-IV datasets. The results are visualized in Fig. 5, where the bars represent classification accuracy (left axis) and the dotted lines denote the number of model parameters (right axis). Observing the bar charts, SSDA consistently yields suboptimal performance. This effectively verifies our hypothesis that indiscriminately merging diverse sources ignores internal distribution shifts, leading to negative transfer. In contrast, FSA-Net achieves recognition accuracies comparable to the MSDA paradigm. This indicates that our method successfully identifies and utilizes the most beneficial source domains while mitigating the noise introduced by irrelevant ones. The advantages of FSA-Net are more pronounced when considering computational cost. As shown by the dotted lines, SSDA and MSDA maintain constant parameter counts at the lowest and highest levels, respectively. Crucially, the complexity of FSA-Net is dynamic, as it adapts to the distributional characteristics of source domain data. As evident in Fig. 5, our model maintains a parameter scale significantly lower than that of MSDA while matching MSDA's performance. This analysis demonstrates that FSA-Net achieves an optimal trade-off. By flexibly reconfiguring source domains, it eliminates the computational redundancy of multi-source models without sacrificing the generalization ability, thereby providing a highly efficient solution for practical EEG emotion recognition systems.

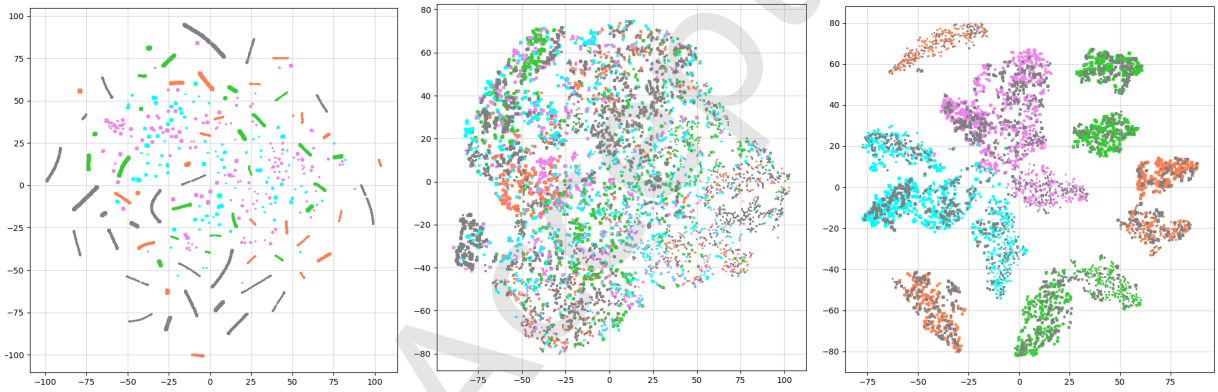
(2) Visualization of feature distributions. To intuitively verify the efficacy of the proposed coarse-to-fine adaptation, we utilize t-SNE [60] to visualize the evolution of feature distributions. We select Subject 6 from the SEED dataset and Subject 1 from the SEED-IV dataset as representative target domains. For clarity, 180 samples are randomly sampled from each aggregated source domain. The visualization results are presented in Fig. 6 and Fig. 7, where markers denote emotion categories, gray points represent the target domain, and colored points represent distinct aggregated source domains. As illustrated in Fig. 6(a) and Fig. 7(a), the raw DE features exhibit severe distribution shifts. Samples from different domains are disjoint, and emotion categories are heavily entangled. This chaotic distribution confirms the inherent non-stationarity of EEG signals and the difficulty of direct transfer. After passing through the common encoder, a preliminary structure emerges. As shown in Fig. 6(b) and Fig. 7(b), samples begin to cluster by emotion category. However, the domain gap remains substantial, indicating that the common encoder extracts general semantic features but lacks sufficient domain invariance. The effectiveness of FSA-Net is fully realized in the private feature space. As revealed in Fig. 6(c) and Fig. 7(c), samples from the target domain (gray) are tightly aligned with the corresponding emotion clusters of the aggregated source domains. The distinct boundaries between emotion categories and the overlap of domain distributions demonstrate that the hierarchical alignment strategy, combining global $\mathcal{L}_{domain}$ and local $\mathcal{L}_{sub}$, has successfully mitigated the domain shift while preserving discriminative class information. The transition from Fig. 6(b) to Fig. 6(c) validates that while the common encoder initiates clustering, the private encoders specialize these features for each source branch, ultimately pulling the target domain into tight alignment with the corresponding emotion clusters. This visualization confirms that FSA-Net progressively refines feature representations, transforming them from a chaotic input space to a domain-invariant, class-discriminative latent space,

Just Accepted



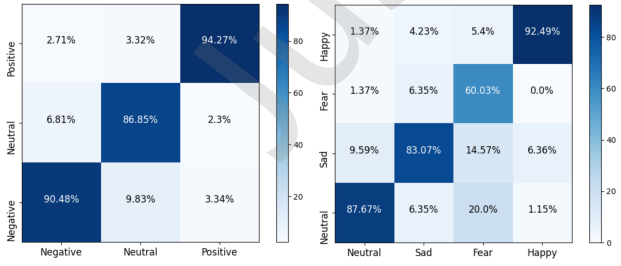
(a) (b) (c)

Fig. 6 The feature visualization of the 6th subject in the SEED dataset with t-SNE. The types of markers $\cdot$, $*$ and $\blacksquare$ denote the different emotion categories: positive, neutral and negative. (a) the original DE feature distributions, the feature distributions after (b) the common encoder, and (c) the private encoder.



(a) (b) (c)

Fig. 7 The feature visualization of the 1st subject in the SEED-IV dataset with t-SNE. The types of markers $\cdot$, $*$, $\blacksquare$ and $+$ denote the different emotion categories: happy, neutral, sad and fear. (a) the original DE feature distributions, the feature distributions after (b) the common encoder, and (c) the private encoder.



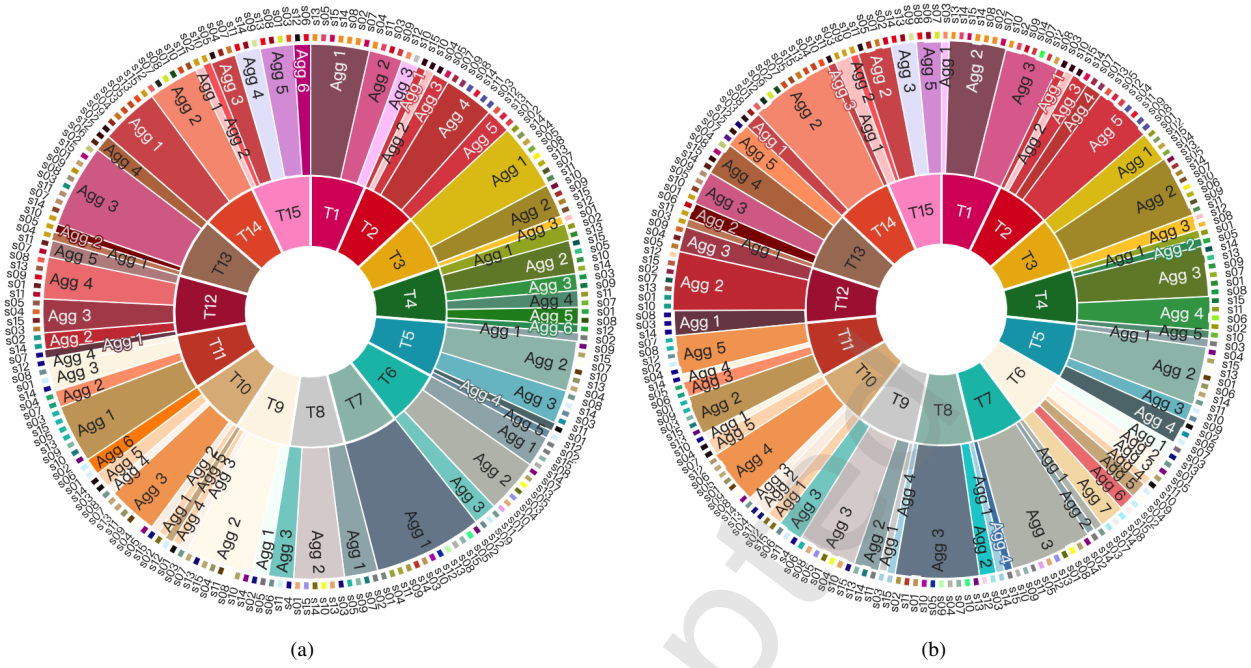
(a) (b)

Fig. 8 The confusion matrices of predictions on the datasets of (a) SEED and (b) SEED-IV.

thereby ensuring robust cross-subject emotion recognition.

(3) Confusion matrix analysis. To further assess the model's sensitivity to specific emotional states, we analyze the confusion matrices on the SEED and SEED-IV datasets, as presented in Fig. 8. On SEED, the FSA-Net achieves impressive recognition accuracies of 94.27% for positive and 90.48% for negative emotions. On SEED-IV, the model accurately identifies 92.49% of happy samples. This confirms that our hierarchical alignment strategy effectively captures the prominent neural signatures associated with strong emotional activations. Conversely, the recognition performance for neutral (on SEED) and fine-grained negative emotions (sad and fear on SEED-IV) is comparatively lower. Notably, for the fear category in SEED-IV, the misclassification rates into neutral and sad reach 20.00% and 14.57%, respectively. From a neurophysiological perspective, the neural activation patterns underlying neutral states and low-arousal negative emotions (such as sadness and fear) share significant spectral similarities and topological overlap. This physiological resemblance creates soft decision boundaries in the latent

Just Accepted



(a) (b)

Fig. 9 The visualization of source domain aggregation on the datasets of (a) SEED and (b) SEED-IV.

feature space, making precise separation inherently more challenging than for distinct emotions like happiness.

(4) Sunburst charts. To provide a comprehensive view of the dynamic aggregation mechanism proposed in FSA-Net, we visualize the specific source configurations for different target subjects using hierarchical sunburst charts, as illustrated in Fig. 9. The innermost circle represents the target subject index, the intermediate ring denotes the resulting aggregated source domains determined by our algorithm, and the outermost layer displays the constituent original source subjects contained within each aggregated group. Crucially, the chart reveals significant variability in the source structures required for different target subjects, empirically validating the necessity of a flexible strategy over rigid paradigms. The algorithm adaptively identifies whether the source distributions are homogeneous or heterogeneous relative to the specific target, thereby optimizing the trade-off between sample diversity and distribution consistency. For instance, taking the 4th subject in the SEED dataset as the target domain, the algorithm identifies high inter-source variance and generates 6 distinct aggregated domains. This configuration preserves fine-grained diversity, resembling an MSDA paradigm. Conversely, for the 7th subject, the algorithm detects high distributional homogeneity among sources and merges them into a single unified domain ($M = 1$), effectively converging to an SSDA paradigm. This adaptive behavior confirms that FSA-Net successfully adjusts the data organization strategy to the unique distributional characteristics of each target subject, avoiding the pitfalls of static approaches.

5 Conclusion

In this work, we presented the flexible source aggregation network (FSA-Net), a unified framework designed to resolve the dilemma between sample diversity and distribution consistency in cross-subject EEG emotion recognition. Unlike rigid single-source or multi-source paradigms, FSA-Net introduces a dynamic similarity-based aggregation mechanism. This strategy effectively bridges the gap between SSDA and MSDA, optimizing the trade-off between model complexity and transfer performance while mitigating negative transfer. To ensure precise feature adaptation, we constructed a hierarchical alignment framework that integrates global discrepancy minimization with local contrastive learning. This coarse-to-fine mechanism not only aligns marginal distributions but also explicitly preserves the discriminative structure of emotion subdomains in the latent space. Furthermore, the proposed domain-aware weighted ensemble strategy enhances the decision-making process by prioritizing high-quality transfer sources. Extensive experiments on the benchmark SEED and SEED-IV datasets demonstrate that FSA-Net achieves state-of-the-art accuracy, validating its efficacy in overcoming individual variability for robust emotion decoding.

Despite these promising results, a limitation lies in the

Just Accepted

determination of the aggregation threshold, which currently relies on empirical sensitivity analysis. Future research will focus on developing adaptive parameter learning mechanisms via meta-learning or reinforcement learning to automatically determine the optimal granularity for source aggregation, further enhancing the model's autonomy and generalization capability.

Acknowledgment

This work is supported by the National Natural Science Foundation of China (Nos. 62501545, 62271455), the Beijing Natural Science Foundation (No. L252143), and the Fundamental Research Funds for the Central Universities (No. CUC26TD03).

Declaration of competing interest

The authors have no competing interests to declare that are relevant to the content of this article.

Reference

[1] I. B. Mauss and M. D. Robinson, "Measures of emotion: a review," *Cognit. Emotion*, vol. 23, no. 2, pp. 209–237, 2010.

[2] A. Luneski, E. Konstantinidis and P. D. Bamidis, "Affective medicine," *Methods Inf. Med.*, vol. 49, no. 3, pp. 207–218, 2010.

[3] B. L. Lu, Y. Q. Zhang and W. L. Zheng, "A survey of affective brain-computer interface," *Chin. J. Intell. Sci. Technol.*, vol. 3, no. 1, pp. 36–48, Mar. 2021.

[4] X. Chen, X. Xu, D. Zhang, and X. Shen, "Exploring the neural basis and validity of ordinal emotion representation through EEG," *Tsinghua Sci. Technol.*, vol. 31, no. 3, pp. 1460–1473, 2026.

[5] J. T. Larsen, G. G. Berntson, K. M. Poehlmann, T. A. Ito and J. T. Cacioppo, "The psychophysiology of emotion," *Handbook of Emotions*, vol. 3, pp. 180–195, 2008.

[6] A. Venkatraman, B. L. Edlow and M. H. Immordino-Yang, "The brainstem in emotion: a review," *Front. Neuroanat.*, vol. 11, no. 15, 2017.

[7] K. Tawsif, N. A. A. Aziz and J. E. Raja, J. Hossen and M. Z. H Jesmeen, "A systematic review on emotion recognition system using physiological signals: data acquisition and methodology," *Emerg. Sci. J.*, vol. 6, no. 5, pp. 1167–1198, 2022.

[8] J. Cheng, M. Chen, C. Li, Y. Liu, R. Song and A. Liu, "Emotion recognition from multi-channel EEG via deep forest," *IEEE J. Biomed. Health Informat.*, vol. 25, no. 2, pp. 453–464, Feb. 2021.

[9] S. Hu, Y. Duan, X. Tao, G. Y. Li, J. Lu, G. Liu, Z. Zheng, and C. Pan, Brain-inspired image perceptual quality assessment based on EEG: A QoE perspective, *IEEE Trans. Pattern Anal. Mach. Intell*., vol. 46, no. 12, pp. 8424–8441, 2024.

[10] S. Hu, Y. Duan, X. Tao, G. Y. Li, and J. Lu, Human perception measurement by electroencephalography for facial image compression, *IEEE Signal Process. Lett*., vol. 29, pp. 2148–2152, 2022.

[11] V. Doma and M. Pirouz, "A comparative analysis of machine learning methods for emotion recognition using EEG and peripheral physiological signals," *J. Big Data*, vol. 7, no. 18, pp. 1–21, 2020.

[12] S. Marjit, U. Talukdar and S. M. Hazarika, "EEG-based emotion recognition using genetic algorithm optimized multi-layer perceptron," in *Proc. Int. Symp. Asian Control Assoc. Intell. Robot. Ind. Autom.*, 2021, pp. 304–309.

[13] Q. Zhang, H. Zhang, K. Zhou, and L. Zhang, "Developing a physiological signal-based, mean threshold and decision-level fusion algorithm (PMD) for emotion recognition," *Tsinghua Sci. Technol.*, vol. 28, no. 4, pp. 673–685, 2023.

[14] Y. Ding, N. Robinson, S. Zhang, Q. Zeng and C. Guan, "Tsception: capturing temporal dynamics and spatial asymmetry from EEG for emotion recognition," *IEEE Trans. Affect. Comput.*, vol. 14, no. 3, pp. 2238–2250, 2023.

[15] V. J. Lawhern, A. J. Solon, N. R. Waytowich, S. M. Gordon, C. P. Hung and B. J. Lance, "EEGNet: a compact convolutional neural network for EEG-based brain-computer interfaces," *J. Neural Eng.*, vol. 15, no. 5, 2018.

[16] Z. Wang, Y. Wang, C. Hu, Z. Yin, Y. Song, "Transformers for EEG-based emotion recognition: a hierarchical spatial information learning model," *IEEE Sensors J.*, vol. 22, no. 5, pp. 4359–4368, 2022.

[17] S. Hu, M. Li, Y. Liu, W. Jiang, B. Geng, W. Zhong, and L. Ye, Audio-visual perceptual quality measurement via multi-perspective spatio-temporal EEG analysis, *Pattern Recognit*., vol. 176, pp. 113156, 2026.

[18] S. Hu, J. Chu, Y. Duan, X. Tao, and J. Lu, Brain-inspired VR video quality assessment based on electroencephalography, in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Cape Town, South Africa, 2024, pp. 169–174.

[19] W. Tao, J. Ye, X. Liu, F. Li, H. Wu, X. Chen, and F. Wan, "SSM-Transformer based dual-scale convolutional neural network for high-performance EEG decoding," *Tsinghua Sci. Technol.*, 2025, doi: 10.26599/TST.2025.9010125.

[20] X. Yu, M. Z. Aziz, M. T. Sadiq, Z. Fan and G. Xiao, "A new framework for automatic detection of motor and mental imagery EEG signals for robust BCI systems," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–12, 2021.

[21] J. Li, S. Qiu, Y. Shen, C. Liu and H. He, "Multisource transfer learning for cross-subject EEG emotion recognition," *IEEE Trans. Cybern.*, vol. 50, no. 7, pp. 3281–3293, 2020.

[22] H. Chen, M. Jin, Z. Li, C. Fan, J. Li and H. He, "MS-MDA: multisource marginal distribution adaptation for cross-subject and cross-session EEG emotion recognition," *Front. Neurosci.*, vol. 15, pp. 778488-778498, 2021.

Just Accepted

[23] M. Gong, W. Zhong, L. Ye, Q Zhang, "MISNet: multi-source information-shared EEG emotion recognition network with two-stream structure," *Front. Neurosci.*, vol. 18, pp. 1293962-1293977, 2024.

[24] Y. P. Lin and T. P. Jung, "Improving EEG-based emotion classification using conditional transfer learning," *Front. Hum. Neurosci.*, vol. 11, 2017.

[25] S. Ran, W. Zhong, F. Hu, L. Ye, and Q. Zhang, Coarse-to-fine domain adaptation for cross-subject EEG emotion recognition with contrastive learning, in *Proc. Chinese Conf. Pattern Recognit. Comput. Vision (PRCV)*, Urumqi, China, 2024, pp. 406–419.

[26] D. Wu, Y. Xu, and B. L. Lu, "Transfer learning for EEG-based brain–computer interfaces: a review of progress made since 2016," *IEEE Trans. Cogn. Develop. Syst.*, vol. 14, no. 1, pp. 4–19, 2020.

[27] J. Wang and Y. Chen, *Introduction to transfer learning: algorithms and practice*. Springer Nature, 2023.

[28] J. Li, S. Qiu, C. Du, Y. Wang and H. He, "Domain adaptation for EEG emotion recognition based on latent representation similarity," *IEEE Trans. Cogn. Develop. Syst.*, vol. 12, no. 2, pp. 344–353, 2019.

[29] X. Du, C. Ma, G. Zhang, J. Li, Y. K. Lai, G. Zhao, X. Deng, Y. J. Liu and H. Wang, "An efficient LSTM network for emotion recognition from multichannel EEG signals," *IEEE Trans. Affect. Comput.*, vol. 13, no. 3, pp. 1528–1540, 2020.

[30] Y. Luo, S. Y. Zhang, W. L. Zheng and B. L. Lu, "WGAN domain adaptation for EEG-based emotion recognition," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2018, pp. 275–286.

[31] W. Li, W. Huan, S. Shao, B. Hou, A. Song, "MS-FRAN: a novel multi-source domain adaptation method for EEG-based emotion recognition," *IEEE J. Biomed. Health Informat.*, vol. 27, no. 11, pp. 5302–5313 2023.

[32] H. Liu, Y. M. Jin, W. L. Zheng, B. L. Lu, "Cross-subject emotion recognition using deep adaptation networks," in *Proc. Int. Conf. Neural Inf. Process.*, Siem Reap, Cambodia, pp. 403–413, 2018.

[33] Y. Ganin, E. Ustinova, H. Ajakan, P. Germain, H. Larochelle, F. Laviolette, M. Marchand, V. Lempitsky and G. Csurka, "Domain-adversarial training of neural networks," *J. Mach. Learn. Res.*, vol. 17, no. 59, pp. 1–35, 2016.

[34] H. Chen, Z. Li, M. Jin and J. Li, "Meernet: multi-source EEG-based emotion recognition network for generalization across subjects and sessions," in *Proc. IEEE Eng. Med. Biol. Soc. (EMBC)*, Mexico, 2021, pp. 6094–6097.

[35] D. Bethge, P. Hallgarten, T. Grosse-Puppendahl, M. Kari, R. Mikut, A. Schmidt and O. Özdenizci, "Domain-invariant representation learning from EEG with private encoders," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, 2022, pp. 1236–1240.

[36] A. Gretton, K. M. Borgwardt, M. J. Rasch, B. Schlkopf, A. J. Smola, "A kernel two-sample test," *J. Mach. Learn. Res.*, vol. 13, no.1, pp. 723–773, 2012.

[37] H. Cui, A. Liu, X. Zhang, X. Chen, J. Liu and X. Chen, "EEG-based subject-independent emotion recognition using gated recurrent unit and minimum class confusion," *IEEE Trans. Affect. Comput.*, vol. 14, no. 4, pp. 2740–2750, 2022.

[38] W. Li, B. Hou, X. Li, Z. Qiu, B. Peng and Y. Tian, "TMLP+ SR-DANN: a domain adaptation method for EEG-based emotion recognition," *Measurement*, vol. 207, pp. 112379, 2023.

[39] Y. An, S. Hu, S. Liu, Z. Gu, Y. Zhang and Y. Zhang, "Spatiotemporal isomorphic cross-brain region interaction network for cross-subject EEG emotion recognition," *Knowl.-Based Syst.*, pp. 114115, 2025.

[40] J. Chang, Z. Zhang, Y. Qian and P. Lin, "Multi-scale hyperbolic contrastive learning for cross-subject EEG emotion recognition," *IEEE Trans. Affect. Comput.*, vol. 16, no. 3, pp. 1716–1731, 2025.

[41] Q. She, C. Zhang, F. Fang, Y. Ma and Y. Zhang, "Multisource associate domain adaptation for cross-subject and cross-session EEG emotion recognition," *IEEE Trans. Instrum. Meas.*, vol. 72, pp. 1–12, 2023.

[42] Y. Chen, X. Xu and X. Qin, "Cross-subject and cross-session EEG emotion recognition based on multisource structural deep clustering," *IEEE Trans. Cogn. Develop. Syst.*, vol. 17, no. 5, pp. 1245–1259, 2025.

[43] S. Hu, Y. Duan, X. Tao, J. Chu, and J. Lu, Brain-inspired visual attention modeling based on EEG for intelligent robotics, *IEEE J. Sel. Top. Signal Process*., vol. 18, no. 3, pp. 431–443, 2024.

[44] S. Hu, Y. Duan, X. Tao, G. Y. Li, and J. Lu, Brain-inspired image quality assessment method based on electroencephalography feature learning, in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Madrid, Spain, 2021, pp. 1–6.

[45] J. Zhu, G. Xu, Q. Hu, B. Wang, T. Zhou and J. Qin, "Dual contrastive training and transferability aware adaptation for multi-source privacy-preserving motor imagery classification," *IEEE Trans. Instrum. Meas.*, vol. 73, pp. 1–13, 2024.

[46] S. Asgarian, Z. Wang, F. Wan, C. M. Wong, F. Liu, Y. Mohsenzadeh, B. Wang and C. X. Ling, "Multi-view contrastive learning for unsupervised domain adaptation in brain-computer interfaces," *IEEE Trans. Instrum. Meas.*, vol. 73, pp. 1–10, 2024.

[47] X. Shen, X. Liu, X. Hu, D. Zhang and S. Song, "Contrastive learning of subject-invariant EEG representations for cross-subject emotion recognition," *IEEE Trans. Affect. Comput.,* vol. 14, no. 3, pp. 2496–2511, 2022.

[48] C. Li, X. Lin, Y. Liu, R. Song, J. Cheng and X. Chen, "EEG-based emotion recognition via efficient convolutional neural network and contrastive learning," *IEEE Sensors J.*, vol. 22, no. 20, pp. 19608–19619, 2022.

[49] Y. Zhu, F. Zhuang, J. Wang, G. Ke, J. Chen, J. Bian, H. Xiong and Q. He, "Deep subdomain adaptation network for image classification," *IEEE Trans. Neural Netw. Learn. Syst.,* vol. 32, no. 4, pp. 1713–1722, 2020.

[50] W. L. Zheng and B. L. Lu, "Investigating critical frequency bands and channels for EEG-based emotion recognition with

Just Accepted

deep neural networks," *IEEE Trans. Auton. Mental Develop.*, vol. 7, no. 3, pp. 162–175, 2015.

[51] W. L. Zheng, W. Liu, Y. Lu, B. L. Lu and A. Cichocki, "Emotionmeter: a multimodal framework for recognizing human emotions," *IEEE Trans. Cybern.*, vol. 49, no. 3, pp. 1110–1122, 2018.

[52] E. Tzeng, J. Hoffman, N. Zhang, K. Saenko, T. Darrell, "Deep domain confusion: maximizing for domain invariance," arXiv preprint arXiv: 1412.3474, 2014.

[53] B. Sun and K. Saenko, "Deep coral: correlation alignment for deep domain adaptation," in *Proc. Comput. Vision Workshops*, Springer, Cham, vol. 9915, 2016, pp. 443–450.

[54] H. Li, Y. M. Jin, W. L. Zheng and B. L. Lu, "Cross-subject emotion recognition using deep adaptation networks," in *Proc. Int. Conf. Neural Inf. Process.*, Springer, Berlin, 2016, pp. 403–413.

[55] Y. Li, L. Wang, W. Zheng, Y. Zong, L. Qi, Z. Cui, T. Zhang and T. Song, "A novel bi-hemispheric discrepancy model for EEG emotion recognition," *IEEE Trans. Cogn. Develop. Syst.*, vol. 13, no. 2, pp. 354–367, 2020.

[56] S. Liu, Z. Wang, Y. An, B. Li, X. Wang and Y Zhang, "DA-CapsNet: a multi-branch capsule network based on adversarial domain adaption for cross-subject EEG emotion recognition," *Knowl.-Based Syst.*, vol. 283, 2024.

[57] Y. An, S. Hu, S. Liu, X. Wang, Z. Gu and Y. Zhang, "LGDAAN-Nets: a local and global domain adversarial attention neural networks for EEG emotion recognition," *Knowl.-Based Syst.*, vol. 318, pp.113613, 2025.

[58] L. Zhu, W. Ding, J. Zhu, P. Xu, Y. Liu, M. Yan and J. Zhang, "Multisource wasserstein adaptation coding network for EEG emotion recognition," *Biomed. Signal Process. Control*, vol. 76, 2022.

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

[60] L. van der Maaten and G. Hinton, "Visualizing data using t-SNE," *J. Mach. Learn. Res.,* vol. 9, no. 86, pp. 2579–2605, 2008.

Author biography



Shuang Ran received the B.E. and M.E. degrees from the Communication University of China, Beijing, in 2013 and 2016, respectively. She is currently pursuing the Ph.D. degree with the Key Laboratory of Media Audio & Video (Communication University of China), Ministry of Education, Beijing, China. Her research interest is affective computing.



Wei Zhong received the B.E. degree in Electronic Engineering and Ph.D. degree in Circuits and Systems from Xidian University in 2004 and 2010, respectively. She is currently a Professor with the Key Laboratory of Media Audio & Video (Communication University of China), Ministry of Education, Beijing, China. Her research focuses on computational neuroscience, affective computing and computer vision.



Shuzhan Hu received the BS degree from the Department of Electronics and Information Engineering, Beihang University, Beijing, China, in 2019, and the PhD degree from the Department of Electronic Engineering, Tsinghua University, Beijing, China, in 2024. He is currently a lecturer with the School of Data Science and Media Intelligence, Communication University of China, Beijing, China. His current research interests include multimedia communications, quality assessment, and cognitive computing.



Long Ye received the B.E. degree in Electronic Engineering from Shandong University, Jinan, China, in 2003, and the M.S. and Ph.D. degrees in Communication and Information Systems from the Communication University of China, Beijing, China, in 2006 and 2012, respectively. He is currently a Professor with the State Key Laboratory of Media Convergence and Communication, Communication University of China, Beijing, China. He is also the Dean of School of Data Science and Media Intelligence, Communication University of China. His research interests include computer vision, multimedia signal processing and virtual reality.



Qin Zhang received the Ph.D. degree in Engineering from The University of British Columbia, Vancouver, BC, Canada, in 1990. He worked as a Research and Development Scientist with the E.E. Department, The University of British Columbia, from 1990 to 1995. In 2004, he served as the Dean of the TCL Industrial Research Institute. He is currently a Professor with the Key Laboratory of Media Audio & Video (Communication University of China), Ministry of Education, Beijing, China. His research interests include computer vision, audio & video processing and virtual reality.

Just Accepted