Robust Federated Hyperspectral Image Clustering

Xiang Yang[1], **Zhengzhong Zhu**[3], **Dayu Hu**[4], **Xiaowen Ma**[5], **Zihao Li**[6], **Wenxuan Tu**[7], **Taichun Zhou**[2], **Wenxin Zhang**[8], **Renxiang Guan**[2*]
[1]South-Central Minzu University, Wuhan, China
[2]National University of Defense Technology, Changsha, China
[3]Sichuan University, Chengdu, China
[4]Northeastern University, Shenyang, China
[5]Zhejiang University, Hangzhou, China
[6]China University of Geosciences, Wuhan, China
[7]Hainan University, Haikou, China
[8]Chinese Academy of Sciences, Beijing, China

xiangyang@mail.scuec.edu.cn, renxiangguan@nudt.edu.cn

Abstract

Hyperspectral image (HSI) clustering facilitates the unsupervised discrimination of complex surface materials but traditionally relies on the idealized assumption of centralized data availability. In real-world scenarios, however, this assumption clashes with data privacy regulations and the physical distribution of data across isolated silos. While federated learning offers a decentralized solution, existing frameworks in remote sensing are predominantly confined to supervised paradigms and struggle to address the heavy reliance on annotations and Non-IID distributions inherent among clients. To overcome these limitations, we propose a novel framework named Robust Federated HSI Clustering(RFHC) that enables collaborative unsupervised learning without requiring raw data exchange. Specifically, we design a dual-encoder architecture that incorporates a Federated Model Weight Augmentation strategy(FMWA), which generates consistent views through local-global network interactions to mitigate the spectral distortion introduced by traditional data augmentation. Furthermore, we develop a hybrid optimization objective that synergizes prototype relationship optimization with contrastive learning. This mechanism utilizes global prototypes to guide local training, effectively stabilizing feature learning against data heterogeneity while ensuring intra-cluster compactness. Extensive experiments on three benchmark HSI datasets demonstrate the effectiveness and superiority of the proposed method against state-of-the-art model (SOTA) federated approaches.

*Corresponding author
https://github.com/XiangYang-Academic/RFHC

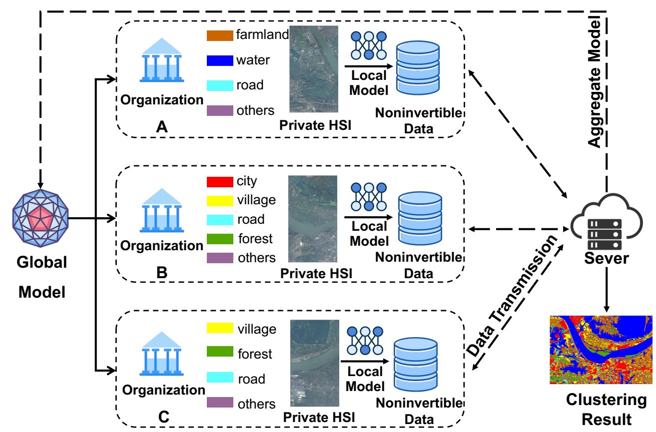

Figure 1: Illustration of the federated HSI clustering paradigm. Distributed clients retain raw data locally to ensure privacy, exchanging only model updates and non-invertible representations with a central server. This enables collaborative multi-source global clustering without transmitting raw data.

1 Introduction

By capturing reflectance across hundreds of narrow bands, a Hyperspectral image (HSI) forms a 3-D tensor that fuses spatial and spectral features [Guan *et al.*, 2024b]. Its near-continuous spectra facilitate precise material discrimination superior to multispectral sensors, effectively identifying subtle land-cover differences. Consequently, HSI is indispensable for tasks such as mineral exploration, urban planning, and resource assessment [Guo *et al.*, 2026; Zhou *et al.*, 2024; Zhao *et al.*, 2025].

To effectively identify and segment land features from massive complex HSI without manual annotation, HSI clustering has emerged as a fundamental unsupervised technique [Zhang *et al.*, 2025]. While shallow methods offer mathematical interpretability, they often lack the capacity to model complex scenes [Cui *et al.*, 2024]. Consequently, deep learning approaches have gained prominence, evolving into three paradigms: multi-stage methods that decouple representation from clustering [Ding *et al.*, 2025], iterative methods that use alternating optimization [Tu *et al.*, 2024], and simultaneous methods that perform end-to-end training [Luo *et al.*, 2024]. These paradigms excel in capturing nonlinear spectral-spatial dependencies, establishing themselves as the dominant methodology for unsupervised HSI analysis.

While existing HSI clustering methods predominantly rely on the idealized assumption of centralized data availability, practical deployment is severely constrained by data silos stemming from strict privacy regulations and communication bottlenecks [Ben Youssef *et al.*, 2024; Büyüktas *et al.*, 2024; Xie *et al.*, 2024]. To overcome these limitations, Federated Learning (FL) has emerged as a transformative solution, facilitating collaborative training across distributed clients without the need for raw data exchange. Nevertheless, existing Federated Remote Sensing (FRS) frameworks remain predominantly confined to supervised paradigms [Gao *et al.*, 2024; Xie *et al.*, 2024], which present two major hurdles for practical deployment. First, the prohibitive cost of pixel-level annotation for HSI necessitates a shift toward unsupervised clustering to reduce reliance on labeled data [Guan *et al.*, 2026; Wang *et al.*, 2025a]. Second, current methods often fail to address the pronounced Non-IID nature of HSI data, leading to model divergence or negative transfer during aggregation. Consequently, there is a critical need for a robust, unsupervised federated HSI clustering framework that can effectively manage data heterogeneity.

To address the aforementioned limitations, we first propose a Robust Federated Hyperspectral Image Clustering (RFHC) framework. As illustrated in Fig. 1, this framework enables collaborative training without sharing private HSI data. Specifically, we propose a dual-encoder architecture augmented with a Federated Model Weight Augmentation(FMWA) strategy. This design generates two distinct views using a locally dedicated encoder and a globally shared encoder to construct positive sample pairs, thereby avoiding the spectral distortion often introduced by traditional data augmentation. It incorporates KAN [Yu *et al.*, 2024] to enhance nonlinear feature extraction, capture complex spectral couplings, and bolster robustness against Non-IID heterogeneity while preserving privacy. To achieve stable training on heterogeneous data, we propose a hybrid objective that integrates contrastive learning with a prototype relationship optimization mechanism. This method integrates contrastive learning with a prototype relationship optimization mechanism. Specifically, the contrastive component enhances the separability and uniformity of representations. Simultaneously, global prototypes guide local training to mitigate the heterogeneity of Non-IID data. Consequently, this synergy stabilizes the training process and facilitates the learning of discriminative semantic representations. To improve HSI scalability, we strategically offload the heavy feature learning module to clients while retaining the lightweight clustering module on the server. In summary, our work makes the following main contributions:

1. **New Research Task:** We pioneer the task of unsupervised federated clustering for HSI. This method enables multi-source collaboration without exchanging raw data.
2. **Novel FHSIC Framework:** We propose FHSIC, a scalable federated clustering framework. It combines customized client-side representation learning with a lightweight server-side clustering module. To address Non-IID data heterogeneity, we introduce prototype relationship optimization and contrastive learning. These mechanisms generate cluster-friendly representations and effectively prevent model collapse.
3. **Better Clustering Results:** We validated our method across three public HSI datasets. The results show superior performance, setting a new SOTA for unsupervised federated clustering.

2 Related Work

2.1 HSI Clustering

HSI clustering partitions pixels into coherent groups using spectral-spatial information without supervision. Existing methods generally fall into shallow and deep categories. Shallow models, utilizing handcrafted features or graph structures [Cui *et al.*, 2024; Zhang *et al.*, 2025; Wang *et al.*, 2025b], often suffer from high computational complexity that limits scalability [Guan *et al.*, 2025a]. Deep clustering methods utilize neural networks through multi-stage, iterative, or simultaneous paradigms [Zhou *et al.*, 2024]. Multi-stage deep clustering refers to a method in which the representation module and the clustering module are optimized separately and then sequentially connected [Guo *et al.*, 2025; Zhao *et al.*, 2025], but risks diminished feature discriminability due to disjoint optimization [Ding *et al.*, 2025]. Iterative deep clustering optimizes the model by alternately executing the clustering module and representation module steps within the training loop [Guan *et al.*, 2024b], yet often incurs high overhead and error propagation risks [Qi *et al.*, 2024; Guan *et al.*, 2025c]. Simultaneous deep clustering optimizes the representation learning module and clustering module end-to-end in parallel [Cai *et al.*, 2021; Cai *et al.*, 2023] but are prone to model collapse without robust regularization [Luo *et al.*, 2024; Guan *et al.*, 2025d]. Despite these advances, the robustness to heterogeneity in federated learning for HSI clustering remains largely unexplored, as most existing methods rely on centralized training.

2.2 Federated Learning for Remote Sensing

FL has emerged as a pivotal distributed paradigm in RS, enabling collaborative model training across data holders without exchanging raw local data to mitigate privacy risks [Moreno-Álvarez *et al.*, 2024]. Although FL facilitates the integration of multi-source data via iterative local-global updates [Ben Youssef *et al.*, 2024], its deployment is hampered by severe data heterogeneity and communication bottlenecks [Büyüktaş *et al.*, 2025]. Existing FRS research predominantly targets supervised tasks, employing various strategies to address these challenges. For instance, to tackle data scarcity and heterogeneity in classification, recent works

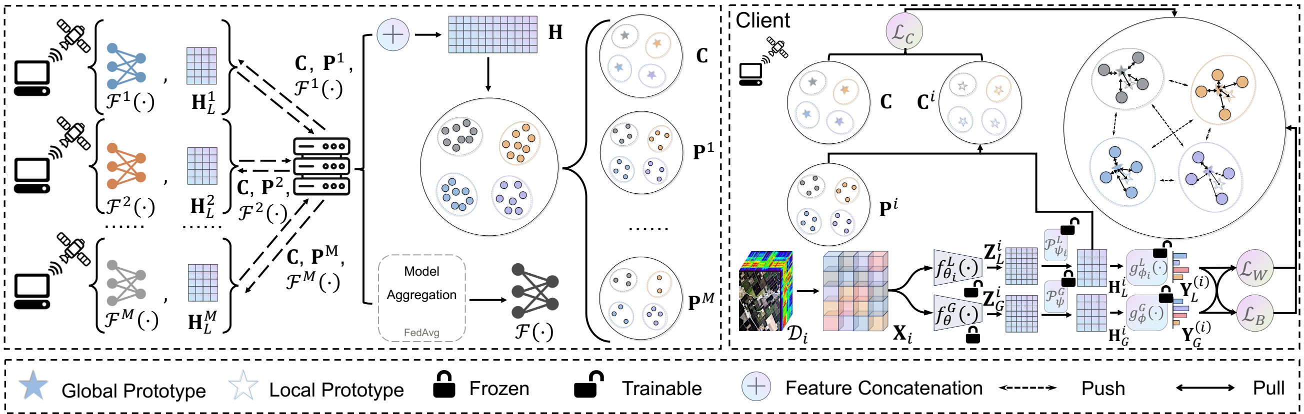

Figure 2: Overview of the proposed RFHC framework. The central server aggregates local parameters and features to update the global model and generate global prototypes. Distributed clients employ a dual-encoder architecture, comprising a trainable local network and a frozen global network, to extract features. A hybrid optimization objective guides local training by aligning local prototypes with global consensus while ensuring instance-level view consistency and inter-cluster separability.

have integrated diffusion models [Li *et al.*, 2024a] or combined CNNs with GCNs [Gao *et al.*, 2024]. In object extraction and segmentation, researchers have introduced prototype matching [Zhang *et al.*, 2023] and geographic heterogeneity-aware frameworks to align class distributions. Additionally, optimization techniques ranging from masked autoencoders to blockchain have been proposed to enhance efficiency and security [Büyüktas *et al.*, 2024]. Despite these advancements, current frameworks exhibit critical limitations for broad-scale HSI analysis: they remain confined to supervised paradigms dependent on unavailable pixel-level annotations, and while addressing general statistical shifts, they fail to resolve the severe Non-IID distributions inherent to decentralized HSI [Moreno-Álvarez *et al.*, 2024; Ben Youssef *et al.*, 2024]. Consequently, the potential of FL for unsupervised HSI clustering remains largely unexplored, necessitating robust frameworks to bridge this gap.

3 Methods

3.1 Problem Definition

We formulate the unsupervised federated HSI clustering task within a Non-IID cross-silo paradigm. Consider a federated system comprising a central server and M clients $\{S_1, \ldots, S_M\}$. Each client S_i maintains a private local HSI dataset $\mathcal{D}_i \in \mathbb{R}^{W^i \times H^i \times D}$ containing a local class set $\mathcal{U}_i$ of size k_i, which exhibits significant heterogeneity in both spatial dimensions and spectral distributions compared with other clients' data [Qi *et al.*, 2024]. The primary objective is to partition the global logical dataset $\mathcal{D} = \bigcup_{i=1}^{M} \mathcal{D}_i$ into $K = \left|\bigcup_{i=1}^{M} \mathcal{U}_i\right|$ semantic clusters without transmitting raw data. We assume a synchronous environment where all clients participate in each communication round [McMahan *et al.*, 2017]. To address this, we introduce the RFHC framework as shown in Figure 2. Let $\mathcal{F}^i(\cdot)$ and $\mathcal{F}(\cdot)$ denote the local network at client S_i and the global network, parameterized by $\{f^L_{\theta_i}, g^L_{\phi_i}, \mathcal{P}^L_{\psi_i}\}$ and $\{f^G_{\theta}, g^G_{\phi}, \mathcal{P}^G_{\psi}\}$, respectively. Consequently, the problem is modeled as a distributed optimization process characterized by two coupled objectives:

Client-Side Objective. Each client S_i aims to extract discriminative high-level features by optimizing parameters $(\theta_i^*, \phi_i^*, \psi_i^*)$ for its personalized local network $\mathcal{F}^i(\cdot)$. This is achieved by minimizing a composite loss function that balances local feature quality with global consistency:

$$\mathcal{L} = \sum_{i=1}^{M} \frac{1}{M} \min \mathcal{L}^i. \tag{1}$$

Server-Side Objective. The server's objective is to aggregate client models to construct a global model $\mathcal{F}(\cdot)$ and to concatenate local embeddings to derive global prototypes $\mathbf{C}$ and client-specific pseudo-labels $\{\mathbf{P}^1, \ldots, \mathbf{P}^M\}$. Ultimately, the server retains the parameters $(\theta^*, \phi^*, \psi^*)$ of $\mathcal{F}(\cdot)$ that yield optimal clustering performance on the global representation $\mathbf{H}$ over T training rounds.

3.2 Local Training

The local training phase optimizes client-specific models to extract discriminative features from private Non-IID data while ensuring global semantic alignment. We propose a specialized architecture that synergizes a dual-encoder specifically for HSI with a prototype-based objective to mitigate model divergence caused by statistical heterogeneity.

Dual Autoencoders Module While many studies employ data augmentation to create positive sample pairs, the rich information inherent in HSI means such transformations risk altering spectral features and corrupting semantic content [Guan *et al.*, 2024b]. Concurrently, the standard autoencoder's reconstruction objective is misaligned to learn discriminative features, as it encourages the latent space to retain excessive input information, creating vulnerabilities for data leakage and impeding the learning of a unified feature representation in a decentralized setting [Balestriero and LeCun, 2024]. To mitigate the curse of dimensionality and spectral distortion from standard augmentation, we introduce Federated Model Weight Augmentation (FMWA). This strategy constructs consistent multi-views via network interaction rather than data transformation. Formally, given pixel cubes $\mathbf{X}_i = \{\mathbf{X}_i^j \in \mathbb{R}^{w\times w\times d}\}_{j=1}^{N^i}$ derived from principal component analysis and sliding windows, FMWA employs a trainable local encoder $f_{\theta_i}^L(\mathbf{X}_i;\theta_i) : \mathbf{X}_i \in \mathbb{R}^{N^i\times d} \longmapsto \mathbf{Z}_L^i \in \mathbb{R}^{N^i\times\hat{d}}$ adapted to private distributions and a frozen global encoder $f_\theta^G(\mathbf{X}_i;\theta) : \mathbf{X}_i \in \mathbb{R}^{N^i\times d} \longmapsto \mathbf{Z}_G^i \in \mathbb{R}^{N^i\times\hat{d}}$ aggregated from collective knowledge. To refine and unify latent representations, we utilize KAN-based projectors $\mathcal{P}_{\psi_i}^L$ and $\mathcal{P}_\psi^G$ to map latent embeddings into high-level representations $\mathbf{H}_L^i$ and $\mathbf{H}_G^i$ [Yu *et al.*, 2024; Guan *et al.*, 2025b]:

$$\mathbf{H}_L^i = \left\{\mathbf{H}_{L,b}^i\right\}_{b=1}^B \triangleq \mathcal{P}_{\psi_i}^L\left(f_{\theta_i}^L(\mathbf{X}_i;\theta_i);\psi_i\right) \in \mathbb{R}^{N^i\times\tilde{d}}, \quad (2)$$

$$\mathbf{H}_G^i = \left\{\mathbf{H}_{G,b}^i\right\}_{b=1}^B \triangleq \mathcal{P}_\psi^G\left(f_\theta^G(\mathbf{X}_i;\theta);\psi\right) \in \mathbb{R}^{N^i\times\tilde{d}}, \quad (3)$$

where $\tilde{d} < \hat{d}$ with $\tilde{d}$ denoting the dimensionality of the high-level representation and B denoting the batch size. This design enables the local network $\mathcal{F}^i(\cdot)$ to extract features invariant to local perturbations yet aligned with the global semantic structure [Li *et al.*, 2024b; Balestriero and LeCun, 2024].

Prototype Relationship Optimization Module Although the dual encoders module provides a robust feature learning foundation, it lacks a mechanism to enforce a consistent global clustering structure. Since the distributed HSI is Non-IID and suffers from the problem of the same object having different spectra, it may erroneously separate pixels of the same type originating from different clients [Kumar and Lee, 2025; Tan *et al.*, 2022]. This forces the model to learn spurious, client-specific features, thereby exacerbating model divergence [Mu *et al.*, 2023; Guan *et al.*, 2025c]. To rectify pixel-level category misalignment arising from Non-IID distributions, this module enforces intra-cluster compactness and inter-cluster separability via prototype-level constraints. In round t, client S_i receives global prototypes $\mathbf{C} = \{\mathbf{c}_1,\ldots,\mathbf{c}_K\} \in \mathbb{R}^{K\times\tilde{d}}$ and pseudo-labels $\mathbf{P}^i \in \mathbb{R}^{N^i}$ where $p_j^i \in \{1,\ldots,K\}$ [Kumar and Lee, 2025; Tan *et al.*, 2022]. We construct an assignment matrix $\mathbf{W}^i \in \mathbb{R}^{K\times N^i}$ where element w_{kj}^i normalizes the contribution of pixel j to cluster k inversely to cluster size [Mu *et al.*, 2023; Guan *et al.*, 2025c]. Local prototypes $\mathbf{C}^i$ are derived by aggregating local features $\mathbf{H}_L^i$ via $\mathbf{W}^i$, allowing gradient backpropagation to the encoder. The k-th local prototype $\mathbf{C}_k^i$ is computed using feature vector $\mathbf{h}_{L,j}^i$ for the j-th pixel as:

$$\mathbf{C}_k^i = \frac{\sum_{j=1}^{N^i} w_{kj}^i \mathbf{h}_{L,j}^i}{\left\|\sum_{j=1}^{N^i} w_{kj}^i \mathbf{h}_{L,j}^i\right\|_2}. \quad (4)$$

To align local geometry with global structure, we minimize the prototype relationship optimization Loss $\mathcal{L}_C^i$. This objective operates on positive pairs formed by corresponding prototypes $\mathbf{C}_k^i$ and $\mathbf{C}_k$ to encourage consistency and negative pairs involving $\mathbf{C}_k^i$ and $\mathbf{C}_j$ for $j \neq k$ to enforce separation. By employing a sampling operation $\mathcal{T}(\cdot)$ on non-corresponding global prototypes $\{\mathbf{C}_j\}_{j\neq k}$ [Tu *et al.*, 2024], the formulation comprises a repulsion term minimizing similarity with negative samples and an attraction term maximizing alignment with positive targets:

$$\mathcal{L}_C^i = \frac{1}{K}\left(\sum_{k=1}^K \frac{\mathbf{C}_k^i\cdot\mathcal{T}(\{\mathbf{C}_j\}_{j\neq k})}{\|\mathbf{C}_k^i\|\,\|\mathcal{T}(\{\mathbf{C}_j\}_{j\neq k})\|} - \sum_{k=1}^K \frac{\mathbf{C}_k^i\cdot\mathbf{C}_k}{\|\mathbf{C}_k^i\|\,\|\mathbf{C}_k\|}\right). \quad (5)$$

Minimizing Eq. (5) guides local features $\mathbf{H}_L^i$ toward global semantic alignment, mitigates the Non-IID problem, and promotes a coherent and globally consistent clustering structure as well as robust feature extraction capability.

3.3 Objective Function

The optimization of the local network $\mathcal{F}^i(\cdot)$ aims to learn personalized features that are structurally aligned with global semantics. The dual-encoder architecture projects mini-batch features $\mathbf{H}_{L,b}^i$ and $\mathbf{H}_{G,b}^i$ into a K-dimensional label space via adaptive local head $g_{\phi_i}^L$ and frozen global head g_ϕ^G, yielding assignment matrices $\mathbf{Y}_{L,b}^i$ and $\mathbf{Y}_{G,b}^i \in \mathbb{R}^{B\times K}$, respectively. The total objective for client S_i is a weighted synergy of structural and instance-level constraints:

$$\mathcal{L}^i = \mathcal{L}_C^i + \alpha_1\mathcal{L}_\mathcal{W}^i + \alpha_2\mathcal{L}_\mathcal{B}^i, \quad (6)$$

where $\alpha_{1,2}$ are balancing hyperparameters. This composite objective establishes a hierarchical optimization strategy: $\mathcal{L}_C^i$ acts as a high-level structural guide, utilizing global prototypes as semantic anchors to alleviate the Non-IID problem. Complementing this, the contrastive duo ($\mathcal{L}_\mathcal{W}^i, \mathcal{L}_\mathcal{B}^i$) performs fine-grained refinement to ensure representational quality and prevent collapse. This hierarchy ensures that personalized features remain firmly aligned with the global consensus, effectively mitigating model divergence [Cai *et al.*, 2023].

Within-Cluster Contrastive Loss ($\mathcal{L}_\mathcal{W}$) To enforce view consistency at the instance level, we employ a symmetric InfoNCE objective [Cai *et al.*, 2023; Cai *et al.*, 2021]. Let $\mathbf{y}_{L,j}^i$ and $\mathbf{y}_{G,j}^i$ denote the j-th row vectors of assignment matrices $\mathbf{Y}_{L,b}^i$ and $\mathbf{Y}_{G,b}^i$, representing the probability distributions of the j-th sample. Treating $(\mathbf{y}_{L,j}^i, \mathbf{y}_{G,j}^i)$ as a positive pair, the loss component anchored on the local view is:

$$\mathcal{L}_{\mathcal{W},L}^{i,j} = -\log\frac{\exp(\mathrm{sim}(\mathbf{y}_{L,j}^i,\mathbf{y}_{G,j}^i)/\tau)}{\sum\limits_{k=1,k\neq j}^{B}\exp(\mathrm{sim}(\mathbf{y}_{L,j}^i,\mathbf{y}_{G,k}^i)/\tau) + \sum\limits_{k=1}^{B}\exp(\mathrm{sim}(\mathbf{y}_{L,j}^i,\mathbf{y}_{L,k}^i)/\tau)}, \quad (7)$$

where $\mathrm{sim}(\mathbf{u},\mathbf{v}) = \mathbf{u}^\top\mathbf{v}/(\|\mathbf{u}\|\|\mathbf{v}\|)$ is the cosine similarity and τ is a temperature hyperparameter. The total loss $\mathcal{L}_\mathcal{W}^i$ enforces symmetry by averaging the objective across both views for all samples in the batch:

$$\mathcal{L}_\mathcal{W}^i = \frac{1}{2B}\sum_{j=1}^{B}\left(\mathcal{L}_{\mathcal{W},L}^{i,j} + \mathcal{L}_{\mathcal{W},G}^{i,j}\right). \quad (8)$$

Between-Cluster Contrastive Loss ($\mathcal{L}_\mathcal{B}$) To enhance inter-cluster separability, we adapt the Barlow Twins objective to minimize redundancy in the label space [Zbontar *et al.*, 2021; Cai *et al.*, 2021]. Distinct from the instance-wise operation in $\mathcal{L}_{\mathcal{W}}$, we focus here on the column vectors $\mathbf{y}^i_{L,k}$ and $\mathbf{y}^i_{G,j}$ of the assignment matrices, which represent the distribution of the k-th cluster across the batch. We compute the cross-correlation matrix $\mathcal{R} \in \mathbb{R}^{K\times K}$, where element $\mathcal{R}_{kj}$ is the cosine similarity between the centered local cluster vector $\mathbf{y}^i_{L,k}$ and global cluster vector $\mathbf{y}^i_{G,j}$:

$$\mathcal{R}_{kj} = \frac{(\mathbf{y}^i_{L,k})^\top \mathbf{y}^i_{G,j}}{\|\mathbf{y}^i_{L,k}\| \|\mathbf{y}^i_{G,j}\|}. \tag{9}$$

The final objective penalizes deviations from the identity matrix:

$$\mathcal{L}^i_{\mathcal{B}} = \sum_{k=1}^{K} (1 - \mathcal{R}_{kk})^2 + \lambda \sum_{k=1}^{K} \sum_{j\neq k} (\mathcal{R}_{kj})^2, \tag{10}$$

where λ is a weighting coefficient controlling the importance of the off-diagonal terms. This formulation comprises an invariance term pushing on-diagonal elements to 1 for alignment and a redundancy reduction term pushing off-diagonal elements to 0 for decorrelation, preventing trivial solutions where all samples collapse to a single cluster.

3.4 Global Training

The global training phase, executed on the central server, orchestrates collaborative learning to refine both the global model and the shared clustering structure. By synthesizing a unified model from decentralized Non-IID data, this phase coordinates two core modules: model aggregation and federated HSI clustering, ensuring global semantic consistency across isolated clients.

Model Aggregation Module The primary objective of this module is to consolidate client-side knowledge into a cohesive global model $\mathcal{F}(\cdot)$, thereby preventing catastrophic forgetting and establishing a shared consensus. At the end of each round, clients upload their optimal parameters $(\theta_i^*, \phi_i^*, \psi_i^*)$. The server aggregates these using FedAvg [McMahan *et al.*, 2017], scaling contributions by the relative dataset size N^i/N to update the global parameters:

$$\begin{bmatrix} \theta \\ \phi \\ \psi \end{bmatrix} = \sum_{i=1}^{M} \frac{N^i}{N} \begin{bmatrix} \theta_i^* \\ \phi_i^* \\ \psi_i^* \end{bmatrix}. \tag{11}$$

The updated components $(f^G_\theta, g^G_\phi, \mathcal{P}^G_\psi)$ are distributed back as the frozen global view, anchoring local training to the global optimum to mitigate client drift.

Federated HSI Clustering Module This module refines the global clustering structure by generating updated global prototypes and pseudo-labels, acting as a global arbiter to enforce a unified semantic partitioning of the feature space. To analyze the global distribution without physically centralizing data, the server constructs a global feature matrix $\mathbf{H}$ by concatenating uploaded local representations $\mathbf{H}^i_L \in \mathbb{R}^{N^i\times \tilde{d}}$. This matrix represents the global logical dataset $\mathcal{D}$:

$$\mathbf{H} = [\mathbf{H}^1_L; \mathbf{H}^2_L; \ldots; \mathbf{H}^M_L]^\top \in \mathbb{R}^{N\times\tilde{d}}. \tag{12}$$

This aggregated matrix provides a holistic view of the entire dataset's feature distribution. Subsequently, we employ the K-means algorithm $\mathcal{K}(\cdot)$ on $\mathbf{H}$ to derive the global semantic structure [Zhang *et al.*, 2023; Tan *et al.*, 2022]. The process yields the final cluster assignment vector $\mathbf{P} \in \mathbb{R}^N$ and the global weight matrix $\mathbf{W} \in \mathbb{R}^{K\times N}$:

$$\{\mathbf{W}, \mathbf{P}\} \leftarrow \mathcal{K}(\mathbf{H}). \tag{13}$$

Derived from these assignments, the updated global prototypes $\mathbf{C} \in \mathbb{R}^{K\times\tilde{d}}$ are constructed to encapsulate the collective semantic consensus. Specifically, the k-th prototype $\mathbf{C}_k$ is synthesized by aggregating the global feature vectors $\mathbf{h}_j$ for the j-th global pixel drawn from the global feature matrix $\mathbf{H}$. The aggregation is performed using weights w_{kj} from the assignment matrix $\mathbf{W}$, which quantify the membership confidence of pixel j to cluster k:

$$\mathbf{C}_k = \frac{\sum_{j=1}^{N} w_{kj}\mathbf{h}_j}{\left\| \sum_{j=1}^{N} w_{kj}\mathbf{h}_j \right\|_2}. \tag{14}$$

This weighted aggregation effectively filters local noise, producing robust prototypes $\mathbf{C}$ and pseudo-labels $\{\mathbf{P}^1, \ldots, \mathbf{P}^M\}$ that are partitioned from $\mathbf{P}$ and guide the subsequent local training toward a globally aligned feature space. Finally, the clustering quality is validated via the nearest-prototype rule:

$$\mathcal{Y} = \left\{ \arg\min_{k\in\{1,\ldots,K\}} \|\mathbf{h}_j - \mathbf{c}_k\|_2^2 \right\}_{j=1}^{N}. \tag{15}$$

4 Experiments

4.1 Experimental Setup

Datasets We evaluate the proposed method on three public hyperspectral datasets: Indian Pines (145×145 pixels, 200 bands, 16 classes), Salinas (512×217 pixels, 204 bands, 16 classes), and Pavia University (610×340 pixels, 103 bands, 9 classes).

Compared Methods To evaluate the performance of our proposed RFHC, we benchmark it against 7 SOTA federated clustering baselines. These include one-shot approaches such as k-FED [Dennis *et al.*, 2021] and NN-FC [Wang *et al.*, 2024], as well as matrix factorization-based algorithms FedMAvg and FedMGS [Wang and Chang, 2022]. Furthermore, we compare RFHC with the cluster-contrastive framework CCFC [Yan *et al.*, 2024] and the lossless fuzzy c-means variants LLF-CM and LLF-FCM [Bárcena *et al.*, 2024].

Implementation Details To ensure fairness, we report averages over ten iterative runs conducted on a 24GB RTX 3090 GPU with 64GB RAM. The training schedule consists of 100 epochs across ten execution runs, utilizing the Adam optimizer with a learning rate of 1×10^{-6} and a weight decay of 5×10^{-4}. The hyperparameters λ, τ, and T are set to 0.005, 0.5, and 100, respectively. Regarding the architecture, we employ a hybrid 3D–2D CNN backbone that consists of stacked convolutional blocks and an MLP projector integrated with a KAN layer to capture complex nonlinear relationships. All

Dataset	Metric(%)	Method							
		k-FED	FedMAvg	FedMGS	CCFC	NN-FC	LLF-CM	LLF-FCM	Ours
Indian Pines	ACC	42.65 ± 2.20	33.48 ± 0.14	30.45 ± 0.15	26.81 ± 1.14	39.48 ± 0.43	36.05 ± 0.83	35.48 ± 3.99	**44.40 ± 2.02**
	Kappa	36.03 ± 2.79	28.46 ± 0.30	24.80 ± 0.19	20.71 ± 0.26	33.24 ± 1.24	29.91 ± 1.10	28.66 ± 3.94	**38.81 ± 2.27**
	NMI	20.20 ± 2.87	34.33 ± 0.19	35.51 ± 0.19	25.36 ± 0.41	44.23 ± 0.75	41.97 ± 1.48	39.97 ± 0.87	**47.48 ± 2.72**
	ARI	**48.49 ± 2.65**	13.86 ± 0.09	12.47 ± 0.03	9.58 ± 0.68	21.55 ± 1.01	19.02 ± 0.81	19.83 ± 3.45	23.38 ± 2.02
Salinas	ACC	50.39 ± 4.78	45.45 ± 0.15	38.35 ± 0.22	31.58 ± 0.32	40.08 ± 0.96	50.49 ± 1.13	49.48 ± 3.61	**60.29 ± 2.99**
	Kappa	43.76 ± 5.55	40.31 ± 0.27	32.29 ± 0.14	19.64 ± 0.27	32.02 ± 1.41	45.62 ± 1.14	42.34 ± 5.15	**56.44 ± 3.29**
	NMI	39.38 ± 4.64	48.63 ± 0.05	48.03 ± 0.13	30.68 ± 0.43	45.93 ± 2.14	51.76 ± 1.09	58.21 ± 4.26	**64.50 ± 2.16**
	ARI	**66.88 ± 1.80**	27.07 ± 0.12	26.08 ± 0.05	13.07 ± 0.84	23.44 ± 1.44	35.82 ± 0.22	33.41 ± 7.11	39.90 ± 1.74
Pavia Univ.	ACC	36.43 ± 2.21	34.20 ± 0.20	32.83 ± 0.13	43.45 ± 0.21	38.35 ± 3.65	37.31 ± 0.61	42.32 ± 2.97	**52.93 ± 2.24**
	Kappa	25.95 ± 4.16	22.70 ± 0.17	22.05 ± 0.08	15.97 ± 3.23	21.34 ± 1.07	31.63 ± 0.92	34.21 ± 2.97	**40.85 ± 2.12**
	NMI	39.53 ± 7.20	24.31 ± 0.09	24.94 ± 0.17	18.19 ± 4.63	24.92 ± 3.70	42.30 ± 0.80	**44.67 ± 1.29**	42.88 ± 1.97
	ARI	16.39 ± 2.38	10.48 ± 0.14	12.13 ± 0.04	16.35 ± 4.08	15.87 ± 4.15	17.91 ± 0.11	22.43 ± 2.69	**33.14 ± 3.16**

Table 1: The clustering results on three HSI datasets. Best and suboptimal results are highlighted in **bold** and underline.

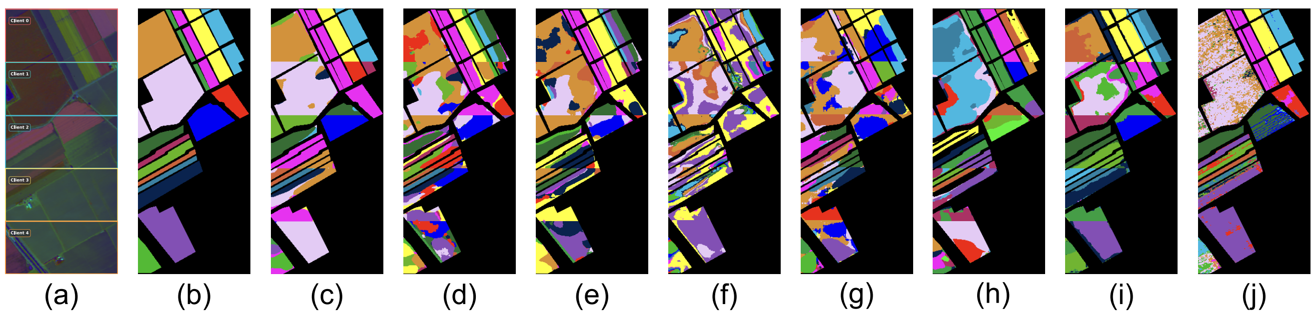

Figure 3: Clustering maps on Salinas dataset. (a) original HSI; (b) ground truth; (c) k-FED; (d) FedMAvg; (e) FedMGS; (f) CCFC; (g) NN-FC; (h) LLF-CM; (i) LLF-FCM; (j) Our proposed RFHC.

baselines adopt their corresponding open-source implementations with optimal parameters. We evaluate performance using overall accuracy (ACC), Kappa, normalized mutual information (NMI), and adjusted Rand index (ARI) [Guan *et al.*, 2024a; Yang *et al.*, 2026; Luo *et al.*, 2024; Cai *et al.*, 2021; Qi *et al.*, 2024], noting that ACC and NMI range within [0, 1] whereas Kappa and ARI range within [-1, 1].

4.2 Experimental Results

We evaluate the clustering performance of RFHC against seven competitive federated baselines across three HSI datasets with 5 clients. The quantitative results, summarized in Table 1, yield the following key observations:

(1) **Overall Superiority:** RFHC consistently achieves dominant performance across the majority of metrics. Particularly on the Pavia University(PU) dataset, RFHC attains an ACC of 52.93% and a Kappa of 40.85%, surpassing the second-best method by 9.48% and the LLF-FCM baseline by 6.64%, respectively. These margins strongly validate the effectiveness of our dual-encoder architecture and prototype optimization in federated settings.

(2) **Comparison with Shallow and Lossless Methods:** While traditional (k-FED) and lossless frameworks (LLF-FCM) exhibit competitive structural metrics, RFHC significantly surpasses them in classification accuracy. Specifically, RFHC outperforms LLF-FCM by margins of 8.92%, 10.81%, and 10.61% in ACC across Indian Pines(IP), Salinas(SA), and PU, respectively. This indicates that KAN-enhanced nonlinear extraction captures complex spectral semantics better than statistical aggregation.

(3) **Comparison with Deep Federated Approaches:** Deep baselines FedMAvg, FedMGS, and CCFC often struggle with Non-IID HSI scenarios. In contrast, RFHC exhibits substantial improvements, notably boosting ACC by 17.59% on IP and 28.71% on SA compared to the contrastive-based CCFC. This confirms that our FMWA strategy and prototype-guided training effectively mitigate model divergence and representation collapse in Non-IID scenarios.

In addition to quantitative metrics, we assess clustering quality visually. As illustrated in Fig. 3, baseline methods exhibit noticeable intra-class imbalance noise and struggle to differentiate spectrally similar agricultural classes. Moreover, they fail to address the ”same object, different spectra” phenomenon, which causes identical categories to be misclassified across different clients. In contrast, RFHC effectively alleviates this heterogeneity to produce smoother and spatially coherent maps with distinct boundaries. This visual superiority is largely attributed to the prototype relationship opti-

Main Loss			Net. Arch.		Indian Pines				Salinas				Pavia University			
$\mathcal{L}_C$	$\mathcal{L}_\mathcal{B}$	$\mathcal{L}_\mathcal{W}$	Opt.	$\mathcal{P}$	ACC(%)	Kappa(%)	ARI(%)	NMI(%)	ACC(%)	Kappa(%)	ARI(%)	NMI(%)	ACC(%)	Kappa(%)	ARI(%)	NMI(%)
			Init.		30.48±2.25	24.86±2.04	31.04±0.92	15.68±1.71	38.54±2.27	32.59±2.44	39.47±3.76	22.03±2.42	32.35±2.53	18.54±1.85	20.54±1.40	9.29±1.74
			Init.	✓	26.20±3.41	18.82±3.26	18.50±1.34	7.79±2.07	37.75±1.92	31.02±2.32	32.69±1.78	19.23±1.12	32.85±1.97	18.57±3.08	16.86±2.13	9.96±1.81
✓	✓	✓	Train		40.79±1.41	35.41±1.18	42.19±2.14	17.80±1.69	56.49±2.91	51.99±3.70	60.39±3.10	38.18±3.33	50.22±2.82	39.00±3.34	42.64±2.70	27.27±3.64
✓			Train	✓	32.20±1.21	24.56±0.95	24.74±1.68	11.23±0.75	50.97±1.71	45.13±1.99	44.17±2.55	30.51±2.87	39.87±1.05	24.16±3.49	30.86±1.86	13.81±2.43
	✓		Train	✓	36.46±2.27	28.70±1.92	36.30±1.91	9.64±1.49	52.01±1.19	46.40±1.09	50.42±2.64	30.91±2.06	42.41±0.61	30.04±2.68	33.90±3.18	15.60±2.70
		✓	Train	✓	35.35±0.73	27.79±1.33	35.20±1.43	9.88±1.79	51.74±1.28	46.80±1.89	54.05±3.44	30.83±3.17	42.24±0.88	28.95±4.84	34.33±3.25	16.75±3.16
✓	✓		Train	✓	35.71±1.24	29.02±1.31	33.33±2.56	11.32±1.52	51.35±1.66	46.01±1.75	48.97±3.36	31.02±3.49	41.80±1.30	26.98±4.24	35.70±2.28	15.89±2.64
✓		✓	Train	✓	38.77±2.35	31.67±1.16	38.06±2.17	11.39±2.20	54.56±3.54	48.98±4.28	51.95±4.21	33.09±5.53	44.94±1.49	32.36±1.42	34.62±2.69	19.13±3.66
	✓	✓	Train	✓	39.06±0.88	33.70±1.24	39.51±4.45	15.93±0.50	54.63±1.23	49.82±1.13	57.68±2.59	32.58±2.64	43.97±3.20	31.89±5.06	37.20±6.52	16.30±6.43
✓	✓	✓	Train	✓	**44.40±2.02**	**38.81±2.27**	**47.48±2.72**	**23.38±2.02**	**60.29±2.99**	**56.44±3.29**	**64.50±2.16**	**39.90±1.74**	**52.93±2.24**	**40.85±2.12**	**42.88±1.97**	**33.14±3.16**

Table 2: Comprehensive ablation study across three datasets. We compare the impact of Main Losses and Network Architecture (Optimization and KAN module $\mathcal{P}$) under identical settings. "Init." refers to Initialization Only. Best results are highlighted in **bold**.

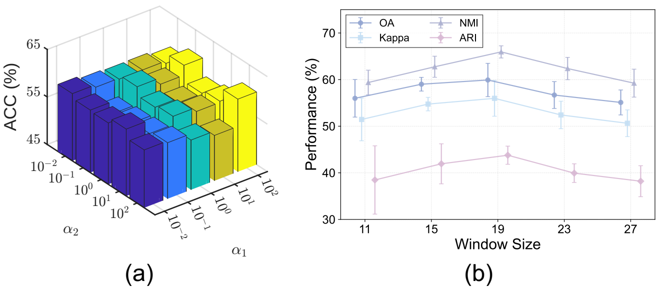

Figure 4: Parameter sensitivity analysis on the SA dataset: (a) Impact of hyper-parameters $\alpha_{1,2}$; (b) Impact of window size w.

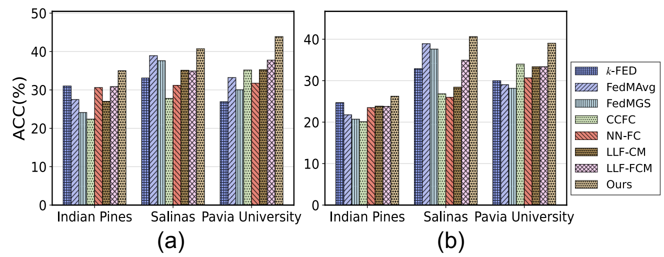

Figure 5: Performance comparison with varying parameter M: (a) $M = 10$; (b) $M = 20$).

mization, which enforces semantic consistency and reduces misclassification in transition regions.

4.3 Ablation Study

Table 2 validates the contribution of each RFHC component.

Ablation on the Main Loss The full objective yields optimal accuracy, demonstrating that these components are mutually reinforcing. Notably, removing the prototype relationship optimization ($\mathcal{L}_C$) causes a sharp decline, underscoring its critical role in aligning local-global representations. Similarly, excluding contrastive losses ($\mathcal{L}_\mathcal{B}, \mathcal{L}_\mathcal{W}$) significantly degrades performance; for instance, removing $\mathcal{L}_\mathcal{W}$ reduces ACC on IP to 35.71%. This confirms that while prototypes handle heterogeneity, contrastive constraints are essential for instance-level discriminability and preventing collapse.

Ablation on the Network Architecture First, the significant gap between random initialization ("Init.") and optimized models ("Train") verifies the framework's learning efficacy. Second, integrating the KAN module ($\mathcal{P}$) enhances representation quality. This demonstrates that KAN effectively captures nonlinear spectral-spatial couplings, thereby bolstering feature robustness against data heterogeneity.

4.4 Parameter Analysis

Impact of Hyper-parameters α_1 and α_2 We examine the regularization weights α_1 and α_2 within $\{10^{-2}, \ldots, 10^2\}$, as visualized in Fig. 4(a). The results reveal that performance degrades significantly when both parameters are small, indicating that the prototype objective alone is insufficient to prevent model collapse. Conversely, increasing these weights enhances accuracy, with the SA dataset achieving optimality at $(\alpha_1, \alpha_2) = (100, 100)$.

Impact of Hyper-parameter w As shown in Fig. 4(b), the SA dataset exhibits a convex trend regarding window size w. Small patches yield suboptimal results due to insufficient spatial context for noise suppression. Performance peaks at $w = 19$ but declines thereafter. Consequently, $w = 19$ is selected to balance spatial smoothing and boundary precision.

Impact of Hyper-parameter M We evaluate scalability by varying the number of clients, where M takes values from the set $\{5, 10, 20\}$, as illustrated in Fig. 5 and Table 1. A consistent performance decline is observed across all datasets as M increases due to intensified data fragmentation; for instance, accuracy on IP decreases from 44.40% when $M = 5$ to 26.23% when $M = 20$.

5 Conclusion

In conclusion, this study presents RFHC, a robust framework for unsupervised federated hyperspectral image clustering that addresses data silos and annotation scarcity. To ensure feature integrity, the model utilizes a dual-encoder architecture with a federated model weight augmentation strategy, generating consistent views without introducing spectral distortion. Furthermore, we incorporate a prototype relationship optimization mechanism combined with KAN layers to align local training with global consensus, effectively mitigating the adverse effects of severe Non-IID distributions.

References

[Balestriero and LeCun, 2024] Randall Balestriero and Yann LeCun. Learning by reconstruction produces uninformative features for perception. *arXiv preprint arXiv:2402.11337*, 2024.

[Bárcena *et al.*, 2024] José Luis Corcuera Bárcena, Francesco Marcelloni, Alessandro Renda, Alessio Bechini, and Pietro Ducange. Federated c-means and fuzzy c-means clustering algorithms for horizontally and vertically partitioned data. *IEEE Transactions on Artificial Intelligence*, 2024.

[Ben Youssef *et al.*, 2024] Belgacem Ben Youssef, Lamyaa Alhmidi, Yakoub Bazi, and Mansour Zuair. Federated learning approach for remote sensing scene classification. *Remote Sensing*, 16(12):2194, 2024.

[Büyüktas *et al.*, 2024] Baris Büyüktas, Gencer Sumbul, and Begüm Demir. Federated learning across decentralized and unshared archives for remote sensing image classification: A review. *IEEE Geoscience and Remote Sensing Magazine*, 2024.

[Büyüktaş *et al.*, 2025] Barış Büyüktaş, Gencer Sumbul, and Begüm Demir. A multi-modal federated learning framework for remote sensing image classification. *arXiv preprint arXiv:2503.10262*, 2025.

[Cai *et al.*, 2021] Yaoming Cai, Zijia Zhang, Yan Liu, Pedram Ghamisi, Kun Li, Xiaobo Liu, and Zhihua Cai. Large-scale hyperspectral image clustering using contrastive learning. *arXiv preprint arXiv:2111.07945*, 2021.

[Cai *et al.*, 2023] Yaoming Cai, Zijia Zhang, Pedram Ghamisi, Behnood Rasti, Xiaobo Liu, and Zhihua Cai. Transformer-based contrastive prototypical clustering for multimodal remote sensing data. *Information Sciences*, 649:119655, 2023.

[Cui *et al.*, 2024] Kangning Cui, Ruoning Li, Sam L Polk, Yinyi Lin, Hongsheng Zhang, James M Murphy, Robert J Plemmons, and Raymond H Chan. Superpixel-based and spatially-regularized diffusion learning for unsupervised hyperspectral image clustering. *IEEE Transactions on Geoscience and Remote Sensing*, 2024.

[Dennis *et al.*, 2021] Don Kurian Dennis, Tian Li, and Virginia Smith. Heterogeneity for the win: One-shot federated clustering. In *International conference on machine learning*, pages 2611–2620. PMLR, 2021.

[Ding *et al.*, 2025] Yao Ding, Zhili Zhang, Aitao Yang, Yaoming Cai, Xiongwu Xiao, Danfeng Hong, and Junsong Yuan. Slcgc: A lightweight self-supervised low-pass contrastive graph clustering network for hyperspectral images. *arXiv preprint arXiv:2502.03497*, 2025.

[Gao *et al.*, 2024] Quanshan Gao, Taixia Wu, and Shudong Wang. Ssc-sfn: spectral-spatial non-local segment federated network for hyperspectral image classification with limited labeled samples. *International Journal of Digital Earth*, 17(1):2300319, 2024.

[Guan *et al.*, 2024a] Renxiang Guan, Zihao Li, Wenxuan Tu, Jun Wang, Yue Liu, Xianju Li, Chang Tang, and Ruyi Feng. Contrastive multi-view subspace clustering of hyperspectral images based on graph convolutional networks. *IEEE Transactions on Geoscience and Remote Sensing*, 2024.

[Guan *et al.*, 2024b] Renxiang Guan, Wenxuan Tu, Zihao Li, Hao Yu, Dayu Hu, Yuzeng Chen, Chang Tang, Qiangqiang Yuan, and Xinwang Liu. Spatial-spectral graph contrastive clustering with hard sample mining for hyperspectral images. *IEEE Transactions on Geoscience and Remote Sensing*, 2024.

[Guan *et al.*, 2025a] Renxiang Guan, Junhong Li, Siwei Wang, Wenxuan Tu, Miaomiao Li, En Zhu, Xinwang Liu, and Ping Chen. Multi-view graph clustering with dual relation optimization for remote sensing data. In *Proceedings of the 33rd ACM International Conference on Multimedia*, pages 7346–7355, 2025.

[Guan *et al.*, 2025b] Renxiang Guan, Junhong Li, Siwei Wang, Wenxuan Tu, Miaomiao Li, En Zhu, Xinwang Liu, and Ping Chen. Multi-view graph clustering with dual relation optimization for remote sensing data. In *Proceedings of the 33rd ACM International Conference on Multimedia*, page 7346–7355, 2025.

[Guan *et al.*, 2025c] Renxiang Guan, Wenxuan Tu, Dayu Hu, Weixuan Liang, Ke Liang, Yaowen Hu, Yue Liu, and Xinwang Liu. Prototype-driven multi-view attribute-missing graph clustering. *IEEE Transactions on Multimedia*, 2025.

[Guan *et al.*, 2025d] Renxiang Guan, Wenxuan Tu, Siwei Wang, Jiyuan Liu, Dayu Hu, Chang Tang, Yu Feng, Junhong Li, Baili Xiao, and Xinwang Liu. Structure-adaptive multi-view graph clustering for remote sensing data. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pages 16933–16941, 2025.

[Guan *et al.*, 2026] Renxiang Guan, Junhong Li, Siwei Wang, Tianrui Liu, Dayu Hu, Miaomiao Li, and Xinwang Liu. Graph masked autoencoder for multi-view remote sensing data clustering. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 40, pages 21396–21404, 2026.

[Guo *et al.*, 2025] Jipeng Guo, Yanfeng Sun, Xin Ma, Junbin Gao, Yongli Hu, Youqing Wang, and Baocai Yin. Globality meets locality: An anchor graph collaborative learning framework for fast multiview subspace clustering. *IEEE Transactions on Neural Networks and Learning Systems*, 36(6):10213 – 10227, 2025.

[Guo *et al.*, 2026] Jipeng Guo, Xiang Xu, Yu Cao, Man Cao, Mengyuan Xin, Tianxiang Zhao, Ye Su, Junbin Gao, Mingliang Cui, and Youqing Wang. Anchor-to-graph structural co-regularization for scalable multi-view clustering. *Pattern Recognition*, 177:113378, 2026.

[Kumar and Lee, 2025] Abhishek Kumar and Dong-Gyu Lee. Enhancing clustering representations with positive proximity and cluster dispersion learning. *Information Sciences*, 686:121277, 2025.

[Li *et al.*, 2024a] Daixun Li, Weiying Xie, Zixuan Wang, Yibing Lu, Yunsong Li, and Leyuan Fang. Feddiff: Diffusion model driven federated learning for multi-modal and multi-clients. *IEEE Transactions on Circuits and Systems for Video Technology*, 2024.

[Li *et al.*, 2024b] Junhong Li, Renxiang Guan, Yuhang Han, Yaowen Hu, Zihao Li, Yanyan Wu, Ziwei Xu, and Xianju Li. Superpixel-based dual-neighborhood contrastive graph autoencoder for deep subspace clustering of hyperspectral image. In *International Conference on Intelligent Computing*, pages 181–192. Springer, 2024.

[Luo *et al.*, 2024] Fulin Luo, Yi Liu, Yule Duan, Tan Guo, Lefei Zhang, and Bo Du. Sdst: Self-supervised double-structure transformer for hyperspectral images clustering. *IEEE Transactions on Geoscience and Remote Sensing*, 2024.

[McMahan *et al.*, 2017] Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, and Blaise Aguera y Arcas. Communication-efficient learning of deep networks from decentralized data. In *Artificial intelligence and statistics*, pages 1273–1282. PMLR, 2017.

[Moreno-Álvarez *et al.*, 2024] Sergio Moreno-Álvarez, Mercedes E Paoletti, Andres J Sanchez-Fernandez, Juan A Rico-Gallego, Lirong Han, and Juan M Haut. Federated learning meets remote sensing. *Expert Systems with Applications*, page 124583, 2024.

[Mu *et al.*, 2023] Xutong Mu, Yulong Shen, Ke Cheng, Xueli Geng, Jiaxuan Fu, Tao Zhang, and Zhiwei Zhang. Fedproc: Prototypical contrastive federated learning on non-iid data. *Future Generation Computer Systems*, 143:93–104, 2023.

[Qi *et al.*, 2024] Jianhan Qi, Yuheng Jia, Hui Liu, and Junhui Hou. Superpixel graph contrastive clustering with semantic-invariant augmentations for hyperspectral images. *IEEE Transactions on Circuits and Systems for Video Technology*, 2024.

[Tan *et al.*, 2022] Yue Tan, Guodong Long, Lu Liu, Tianyi Zhou, Qinghua Lu, Jing Jiang, and Chengqi Zhang. Fedproto: Federated prototype learning across heterogeneous clients. In *Proceedings of the AAAI conference on artificial intelligence*, volume 36, pages 8432–8440, 2022.

[Tu *et al.*, 2024] Wenxuan Tu, Renxiang Guan, Sihang Zhou, Chuan Ma, Xin Peng, Zhiping Cai, Zhe Liu, Jieren Cheng, and Xinwang Liu. Attribute-missing graph clustering network. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pages 15392–15401, 2024.

[Wang and Chang, 2022] Shuai Wang and Tsung-Hui Chang. Federated matrix factorization: Algorithm design and application to data clustering. *IEEE Transactions on Signal Processing*, 70:1625–1640, 2022.

[Wang *et al.*, 2024] Yizhang Wang, Wei Pang, and Witold Pedrycz. One-shot federated clustering based on stable distance relationships. *IEEE Transactions on Industrial Informatics*, 2024.

[Wang *et al.*, 2025a] Xinxin Wang, Yongshan Zhang, and Yicong Zhou. Bidirectional probabilistic multi-graph learning and decomposition for multi-view clustering. *IEEE Transactions on Image Processing*, 34:3609–3621, 2025.

[Wang *et al.*, 2025b] Xinxin Wang, Yongshan Zhang, and Yicong Zhou. Pseudo-supervision affinity propagation for efficient and scalable multiview clustering. *IEEE Transactions on Neural Networks and Learning Systems*, 36(8):15282–15293, 2025.

[Xie *et al.*, 2024] Weiying Xie, Xiaohong Xu, and Yunsong Li. Decentralized federated gan for hyperspectral change detection in edge computing. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 17:8863–8874, 2024.

[Yan *et al.*, 2024] Jie Yan, Jing Liu, and Zhong-Yuan Zhang. Ccfc: Bridging federated clustering and contrastive learning. *arXiv preprint arXiv:2401.06634*, 2024.

[Yang *et al.*, 2026] Jie Yang, Cheng-You Lu, Zhongli Wang, Hsiang-Ting Chen, Guang-Kui Xu, Chenglong Zhang, Shuting Dong, Xinyan Liang, and Bingbing Jiang. Multi-view clustering with granularity-aware pseudo supervision. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 40, pages 27538–27546, 2026.

[Yu *et al.*, 2024] Fangchen Yu, Ruilizhen Hu, Yidong Lin, Yuqi Ma, Zhenghao Huang, and Wenye Li. Kae: Kolmogorov-arnold auto-encoder for representation learning. *arXiv preprint arXiv:2501.00420*, 2024.

[Zbontar *et al.*, 2021] Jure Zbontar, Li Jing, Ishan Misra, Yann LeCun, and Stéphane Deny. Barlow twins: Self-supervised learning via redundancy reduction. In *International conference on machine learning*, pages 12310–12320. PMLR, 2021.

[Zhang *et al.*, 2023] Xiaokang Zhang, Boning Zhang, Weikang Yu, and Xudong Kang. Federated deep learning with prototype matching for object extraction from very-high-resolution remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 61:1–16, 2023.

[Zhang *et al.*, 2025] Zijia Zhang, Yaoming Cai, Wenyin Gong, Xiaobo Liu, Cheng Zeng, and Gan Yu. Mmagl: Multi-objective multi-view attributed graph learning for joint clustering of hyperspectral and lidar data. *IEEE Transactions on Geoscience and Remote Sensing*, 2025.

[Zhao *et al.*, 2025] Yang Zhao, Zixuan Bi, Peican Zhu, Aihong Yuan, and Xuelong Li. Deep spectral clustering with projected adaptive feature selection. *IEEE Transactions on Geoscience and Remote Sensing*, 2025.

[Zhou *et al.*, 2024] Sheng Zhou, Hongjia Xu, Zhuonan Zheng, Jiawei Chen, Zhao Li, Jiajun Bu, Jia Wu, Xin Wang, Wenwu Zhu, and Martin Ester. A comprehensive survey on deep clustering: Taxonomy, challenges, and future directions. *ACM Computing Surveys*, 57(3):1–38, 2024.