Anchor-Guided Discriminative Subspace Alignment and Clustering for Cross-Scene Hyperspectral Imagery

Yongshan Zhang[1], Zixuan Zhang[1], Xinxin Wang[2*], Lefei Zhang[3], Zhihua Cai[1]

[1]School of Computer Science, China University of Geosciences, Wuhan, China
[2]School of Artificial Intelligence, Shenzhen University, Shenzhen, China
[3]School of Computer Science, Wuhan University, Wuhan, China
yszhang.cug@gmail.com, zzixuan@cug.edu.cn, xinxinwang1024@gmail.com, zhanglefei@whu.edu.cn, zhcai@cug.edu.cn

Abstract

Cross-scene hyperspectral image (HSI) recognition aims to assign a unique label to each pixel in the target scene by transferring knowledge from the source scene. Existing methods primarily rely on fully labeled source data and either partially labeled or unlabeled target data. No prior work has addressed the more challenging scenario of cross-scene recognition without label guidance in both scenes. To bridge this gap, we present the first study on cross-scene HSI clustering, proposing an anchor-guided discriminative subspace alignment and clustering (ADSAC) framework that follows a well-structured three-step learning paradigm to effectively mitigate distribution shifts. Specifically, we first develop an anchor-promoted graph learning (APGL) model to efficiently derive accurate clustering labels for the source scene by leveraging anchor-based structural information. Next, we propose a discriminative cross-scene subspace alignment (DCSA) model to improve feature discriminability and reduce distribution discrepancies. Finally, labels of the target scene are inferred after source clustering and cross-scene alignment. To solve the formulated models, we design tailored optimization algorithms to ensure high-quality learning. Extensive experiments demonstrate the superiority of the proposed framework over state-of-the-art methods.

Code — https://github.com/ZhangYongshan/ADSAC

Introduction

Hyperspectral images (HSIs) are captured by specialized sensors across a wide range of continuous spectral bands to record the spatial distribution of the observed area. The rich spatial and spectral information inherent in HSIs makes them highly valuable in diverse applications, such as urban planning, precision agriculture and military defense (Ghamisi et al. 2017). HSI recognition is a fundamental task that enhances the practicality of these applications (Ma et al. 2019; Plaza et al. 2009). A large number of effective methods have been proposed, typically assuming that training and testing data share the same spectral characteristics and spatial distribution (Li et al. 2019; Zhang, Zhang, and Zhou 2023). This assumption is often violated in real-world scenarios. When training and testing data come from different environments or sensors, notable differences in spectral reflectance and spatial distribution can occur, resulting in the distribution shift problem (Chen et al. 2024).

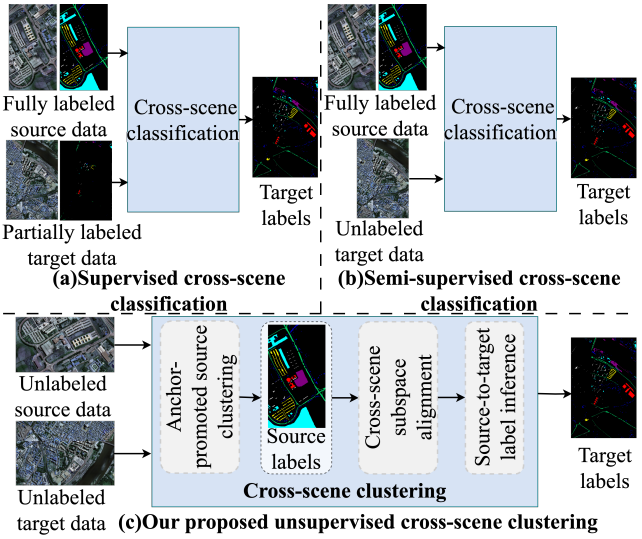

Figure 1: Different cross-scene HSI recognition styles. Our proposed cross-scene clustering method is more challenging without requiring explicit source and target labels.

To address this problem, cross-scene recognition has been explored as a potential solution motivated by the principles of domain adaptation (Kouw and Loog 2019; Sohn et al. 2017; Xu et al. 2022). Given two datasets capturing different scenes but sharing the same land-cover categories, it enables knowledge transfer from the source scene to the target scene without requiring repeated model training. There are numerous effective cross-scene classification methods (Ding et al. 2024; Li et al. 2024b). Depending on label availability, existing methods fall into supervised and semi-supervised categories. As illustrated in Fig. 1 (a), supervised methods perform cross-scene learning using fully labeled source data and partially labeled target data (Li et al. 2024a; Ye et al. 2024). In contrast, as shown in Fig. 1 (b), semi-supervised methods offer greater flexibility by requiring fully labeled source data and unlabeled target data for cross-scene learning (Tang, Li, and Peng 2022; Zhang et al. 2023). Their ma-

*Corresponding author.

jor difference lies in the labeling requirements of the target scene.

Although existing methods show strong generalization performance, they always rely on fully labeled source data and either partially labeled or unlabeled target data. However, in practical scenarios, pixel-level annotation for HSIs is costly and time-consuming (Jiang et al. 2025; Zhang et al. 2024a). The labeling requirement makes these methods inapplicable when explicit labels are unavailable for both source and target data. In such cases, as shown in Fig. 1 (c), cross-scene HSI clustering needs to be explored as a viable alternative. The absence of labels in both scenes significantly increases the difficulty of this task, demanding effective techniques as solutions. To tackle this challenge, three key problems should be considered. (1) Label information is crucial for effective knowledge transfer from source data to target data (Zhang et al. 2021). This leads to the challenge of *how to obtain accurate clustering labels for the source data*. (2) Distribution shift arises from significant discrepancies between two scenes (Xie et al. 2024). This presents the challenge of *how to perform effective subspace alignment to reduce the distribution shift*. (3) Clustering labels of the source data can be served as valuable clues (Zhang et al. 2024b). This highlights the challenge of *how to infer clustering labels for the target data by transferring knowledge from the clustering-labeled source data.*

To solve the above challenges, in this paper, we propose a novel anchor-guided discriminative subspace alignment and clustering (ADSAC) framework for cross-scene HSIs. It follows a well-structured three-step learning process to mitigate distribution shifts and address cross-scene clustering. Specifically, an anchor-promoted graph learning (APGL) model is developed for effective source clustering, while a discriminative cross-scene subspace alignment (DCSA) model is designed to reduce distribution shifts. Target labels are inferred by applying k-nearest neighbors (KNN) trained on clustering-labeled source data and target data within the aligned subspaces. Customized optimization algorithms are devised to solve the formulated models. In summary, the main contributions of this paper are as follows.

- We propose a novel ADSAC framework to achieve cross-scene HSI clustering. To the best of our knowledge, this is the first study on cross-scene HSI recognition without label guidance in both source and target scenes.
- We develop an APGL model to effectively derive clustering labels for the source scene by exploring anchor-based structural information.
- We design a DCSA model to improve feature discriminability and reduce distribution discrepancies between scenes for better cross-scene learning.
- Tailored optimization algorithms are devised to solve the formulated models, and extensive experiments validate the superiority of the proposed framework.

Related Work

Cross-scene HSI recognition is generalized from domain adaption and has attracted increasing attention in recent years (Li et al. 2024a; Xie et al. 2024; Zhao et al. 2022). Supervised and semi-supervised cross-scene classification are the two main tasks in this field. Specifically, Supervised methods predict labels for unlabeled target data by leveraging fully labeled source data and partially labeled target data (Li et al. 2024a). Ye et al. (Ye et al. 2024) proposed a cross-domain discriminative vision transformer to enhance feature alignment and cross-scene classification. Additionally, Li et al. (Li et al. 2024a) introduced a spectral coordinate transformer that employs an intra-domain loss function for robust cross-domain classification. However, when labels are unavailable for target data, these methods are inapplicable. Semi-supervised classification addresses this issue by training a model solely on labeled source data and directly applying it to unlabeled target data (Zhang et al. 2021). Tang et al. (Tang, Li, and Peng 2022) presented an unsupervised joint adversarial domain adaptation architecture that minimizes domain discrepancies through class-level feature alignment. Furthermore, Zhang et al. (Zhang et al. 2023) developed a single-source domain expansion network to enhance domain-invariant feature representation and classification robustness. Nevertheless, when explicit labels are absent for both source and target data, these semi-supervised methods also become ineffective, necessitating alternative strategies.

Unlike existing methods relying on labeled data, unsupervised cross-scene clustering is more flexible by eliminating the need for explicit labels in both scenes, making it well-suited for real-world applications. To address this challenge, we introduce our proposed framework in detail.

Proposed Method

The problem definition is introduced here. The source and target scenes are represented as $\mathbf{X}_s = \{\mathbf{x}_i^s\}_{i=1}^{N_s} \in \mathbb{R}^{D\times N_s}$ with N_s pixels and $\mathbf{X}_t = \{\mathbf{x}_i^t\}_{i=1}^{N_t} \in \mathbb{R}^{D\times N_t}$ with N_t pixels respectively, where each pixel is described by a D-dimensional vector. They exhibit differences in spectral characteristics and spatial distribution. For cross-scene clustering, explicit labels are unavailable for both scenes. This task is defined as $\{\mathbf{Y}_s, \mathbf{Y}_t\} \leftarrow \mathcal{F}(\mathbf{X}_s, \mathbf{X}_t)$, where $\mathcal{F}(\cdot)$ denotes the learning process to reduce distribution discrepancies and obtain clustering labels $\mathbf{Y}_s$ and $\mathbf{Y}_t$ for both scenes. Fig. 2 displays the overall framework of the proposed ADSAC. It consists of three sequential learning steps, including anchor-promoted source clustering, cross-domain subspace alignment and source-to-target label inference.

Anchor-Promoted Source Clustering

In this task, neither the source nor the target data contain explicit labels. The first essential task is to derive meaningful labels for the source data. To this end, we propose an anchor-promoted graph learning (APGL) model. Specifically, to capture the spatial texture, entropy rate superpixel (ERS) segmentation (Liu et al. 2011) is applied to the first principal component of the source scene, enabling the identification of M_s homogeneous regions. Within each generated superpixel, all pixels are averaged to derive a representative anchor, thereby revealing the underlying data distribution. Thus, we can obtain an anchor matrix $\mathbf{A} =$

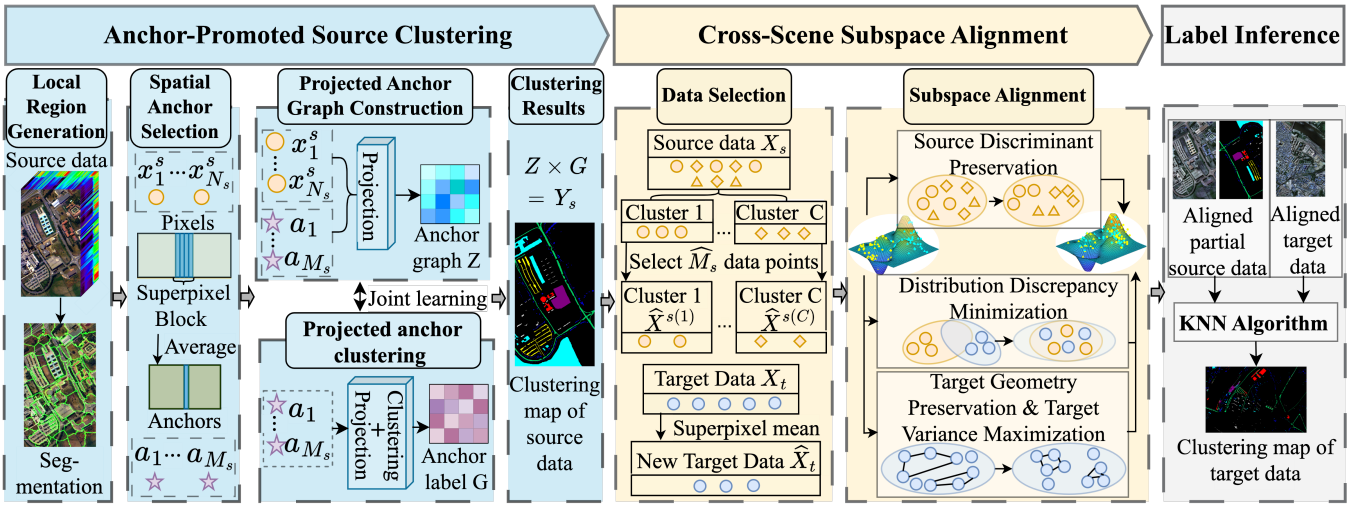

Figure 2: Illustration of our ADSAC. There are three sequential steps to address distribution shifts and cross-scene clustering.

$\{\mathbf{a}_i\}_{i=1}^{M_s} \in \mathbb{R}^{D\times M_s}$, containing M_s high-quality anchors, each with D dimensions rich in spatial information. Based on these, APGL simultaneously perform anchor graph construction and anchor-guided clustering exploration to uncover the clustering structure of both anchors and pixels in the source data. Thus, the proposed APGL is formulated by

$$\begin{aligned}\min_{\mathbf{P},\mathbf{A},\mathbf{Z},\mathbf{F},\mathbf{G}} & \underbrace{\sum_{i=1}^{N_s}\sum_{j=1}^{M_s}\left\|\mathbf{P}^T\mathbf{x}_i^s-\mathbf{P}^T\mathbf{a}_j\right\|_2^2 z_{ij}+\beta\|\mathbf{Z}\|_F^2}_{\text{anchor graph construction}} \\ & + \underbrace{\lambda\left\|\mathbf{P}^T\mathbf{A}-\mathbf{F}\mathbf{G}^T\right\|_F^2}_{\text{anchor-guided clustering exploration}} \\ \text{s.t. } & \mathbf{z}_i^T\mathbf{1}=1,\ z_{ij}\geq 0,\ \mathbf{P}^T\mathbf{X}\mathbf{X}^\mathbf{T}\mathbf{P}=\mathbf{I}, \\ & g_{jk}\in\{0,1\},\ \mathbf{g}_j^T\mathbf{1}=1,\end{aligned} \tag{1}$$

where the first and second terms describe the construction of anchor graph $\mathbf{Z}\in\mathbb{R}^{N_s\times M_s}$ by measuring the distance between pixel $\mathbf{x}_i^s$ and anchor $\mathbf{a}_j$, with their similarity $z_{ij}\in\mathbf{Z}$ being inversely proportional to this distance. By introducing a projection matrix $\mathbf{P}\in\mathbb{R}^{D\times B}$, the anchor graph can be learned from B-dimensional reduced representations, enabling effective structure preservation and reducing redundancy. The third term denotes anchor-guided clustering exploration, where the projected anchors are decomposed into a clustering centroid matrix $\mathbf{F}\in\mathbb{R}^{B\times C}$ and a clustering indicator matrix $\mathbf{G}\in\mathbb{R}^{M_s\times C}$ to reveal their C intrinsic clusters. Appropriate constraints are imposed to ensure effective graph construction and clustering exploration.

The joint learning of anchor graph and cluster distribution effectively uncovers the clustering structure for the source scene. Notably, clustering exploration is performed directly on the anchors rather than the pixels, as the number of anchors is much smaller than that of pixels ($M_s \ll N_s$), leading to efficient learning. After optimizing Eq. (1) for APGL, clustering labels for the anchors are obtained from $\mathbf{G}$. Taking $\mathbf{G}$ as a clue, clustering labels for the source pixels can be calculated as $\mathbf{Y}^s=\mathbf{Z}\mathbf{G}$. This ensures accurate clustering labels for the source data, facilitating subsequent learning.

Cross-Scene Subspace Alignment

Distribution discrepancies between different scenes should be minimized for effective cross-scene learning. This can be achieved through feature alignment within their respective subspaces (Zhang et al. 2021). To enhance efficiency, we perform subspace alignment using representative samples selected from both the source and target data, rather than the entire datasets. Specifically, $\hat{M}_s$ source samples are selected from each clusters based on the results obtained by APGL. They are denoted as $\hat{\mathbf{X}}_s=\bigcup_{c=1}^{C}\{\hat{\mathbf{x}}_i^{s(c)}\}_{i=1}^{\hat{M}_s}\in\mathbb{R}^{D\times\hat{N}_s}$, where $\hat{\mathbf{x}}_i^{s(c)}$ represents the i-th sample selected from cluster c and $\hat{N}_s=C\hat{M}_s$. This sampling strategy effectively preserves discriminative information across different source clusters. Due to the absence of target labels, it is not feasible to select target samples in the same manner as those from the source data. Instead, representative target samples are obtained by independently averaging the pixels within each of the $\hat{N}_t$ superpixels of the target scene. They are denoted as $\hat{\mathbf{X}}_t=\{\hat{\mathbf{x}}_i^t\}_{i=1}^{\hat{N}_t}\in\mathbb{R}^{D\times\hat{N}_t}$, where $\hat{\mathbf{x}}_i^t$ represents as the mean vector of the i-th target superpixel. This sampling strategy leverages the spatial distribution of the target scene, similar to the anchor generation strategy described previously.

Based on the selected samples, we propose a discriminative cross-scene subspace alignment (DCSA) model to learn two projection matrices $\mathbf{W}_s,\mathbf{W}_t\in\mathbb{R}^{D\times R}$ for the source and target scenes, respectively. These matrices enable cross-scene subspace alignment, ensuring that: (1) the discriminative information of clusters in the source scene is preserved, (2) the intrinsic geometric structure of the target scene is maintained, (3) the variance within the target scene is maximized, and (4) the distribution discrepancies between the two scenes are minimized. This alignment facilitates effective knowledge transfer while preserving structural integrity and improving cross-scene adaptability.

Source Discriminant Preservation. To enhance the discriminative capability of source samples, we enforce intra-cluster compactness and inter-class separability. Specifically, samples within the same cluster are encouraged to be tightly grouped, while those from different clusters are pushed apart after subspace alignment (Fisher 1936). Thus, the source discriminant preservation can be achieved by

$$\min_{\mathbf{W}_s} \frac{\mathrm{Tr}(\mathbf{W}_s^T \mathbf{D}_w \mathbf{W}_s)}{\mathrm{Tr}(\mathbf{W}_s^T \mathbf{D}_b \mathbf{W}_s)} \tag{2}$$

where $\mathbf{D}_w$ and $\mathbf{D}_b \in \mathbb{R}^{D\times D}$ are the within-cluster and between-cluster distance matrices. They are defined by

$$\mathbf{D}_w = \sum_{c=1}^{C} \frac{1}{\hat{M}_s^2} \sum_{i=1}^{\hat{M}_s} \sum_{j=1}^{\hat{M}_s} (\hat{\mathbf{x}}_i^{s(c)} - \hat{\mathbf{x}}_j^{s(c)})(\hat{\mathbf{x}}_i^{s(c)} - \hat{\mathbf{x}}_j^{s(c)})^T \tag{3}$$

$$\mathbf{D}_b = \sum_{\substack{c=1 \\ c'\neq c}}^{C} \frac{1}{\hat{M}_s \hat{M}_s'} \sum_{i=1}^{\hat{M}_s} \sum_{j=1}^{\hat{M}_s'} (\hat{\mathbf{x}}_i^{s(c)} - \hat{\mathbf{x}}_j^{s(c')})(\hat{\mathbf{x}}_i^{s(c)} - \hat{\mathbf{x}}_j^{s(c')})^T \tag{4}$$

where $\hat{\mathbf{x}}_i^{s(c)}$ and $\hat{\mathbf{x}}_j^{s(c')}$ represent source samples inside and outside cluster c, respectively. $\hat{M}_s$ and $\hat{M}_s'$ denote the number of samples within and outside cluster c, respectively, such that $\hat{M}_s + \hat{M}_s' = \hat{N}_s$. Unlike traditional methods that assume a Gaussian distribution and identical class covariance matrices, Eq. (2) relaxes these assumptions, allowing for better adaptation to distribution differences among clusters and enhancing both flexibility and adaptability.

Target Geometry Preservation. The geometric structure exhibits an invariant property that is widely utilized in unsupervised learning methods. According to the manifold theorem (Tenenbaum, Silva, and Langford 2000), neighboring samples in the original space are likely to share the same label, and this property remains valid in the subspace. To enforce proximity among similar samples, we formulate the target geometry preservation as the following problem:

$$\begin{aligned} \min_{\mathbf{W}_t} \frac{1}{2} \sum_{i,j=1}^{\hat{N}_t} \|\mathbf{W}_t \hat{\mathbf{x}}_i^t - \mathbf{W}_t \hat{\mathbf{x}}_j^t\|_2^2 b_{ij} \\ = \mathrm{Tr}(\mathbf{W}_t^T \hat{\mathbf{X}}_t \mathbf{L}_t \hat{\mathbf{X}}_t^T \mathbf{W}_t) \end{aligned} \tag{5}$$

where $b_{ij} = \exp(-\frac{\|\hat{\mathbf{x}}_i^t - \hat{\mathbf{x}}_j^t\|^2}{2\sigma^2})$ is an element of graph $\mathbf{B} \in \mathbb{R}^{\hat{N}_t \times \hat{N}_t}$, measuring the similarity between target samples $\hat{\mathbf{x}}_i^t$ and $\hat{\mathbf{x}}_j^t$ based on their p-nearest neighbor relationships. The related graph Laplacian matrix is defined as $\mathbf{L}_t = \mathbf{D} - \mathbf{B}$, where $\mathbf{D} = \mathrm{diag}(\sum_{j=1}^{\hat{N}_t} b_{ij})$.

Target Variance Maximization. To preserve the most informative features while minimizing redundancy, the variance of target samples should be maximized. This can enhance cluster separability within the subspace and mitigate the overlap between distinct clusters. Inspired by principal component analysis (PCA) (Maćkiewicz and Ratajczak 1993), we formulate the target variance maximization by

$$\max_{\mathbf{W}_t} \mathrm{Tr}(\mathbf{W}_t^T \hat{\mathbf{X}}_t \mathbf{H}_t \hat{\mathbf{X}}_t^T \mathbf{W}_t) \tag{6}$$

where $\mathbf{H}_t = \mathbf{I}_t - \frac{1}{\hat{N}_t} \mathbf{1}_t \mathbf{1}_t^T \in \mathbb{R}^{\hat{N}_t \times \hat{N}_t}$ is the centering matrix that ensures the mean of target samples being zero. Here, $\mathbf{I}_t \in \mathbb{R}^{\hat{N}_t \times \hat{N}_t}$ is the identity matrix, and $\mathbf{1}_t \in \mathbb{R}^{\hat{N}_t \times 1}$ is an all-ones column vector.

Distribution Discrepancy Minimization. To achieve subspace alignment, conditional distribution discrepancies between different scenes should be minimized. For HSIs, the conditional distribution denotes the probability distribution of spectral features given a category label. Thus, we use the category-wise mean to model this. As target labels are unavailable, we first train an initial classifier on source samples and generate pseudo-labels for target samples. Based on the maximum mean discrepancy (MMD) (Gretton et al. 2012), the distribution discrepancy minimization is formulated by

$$\min_{\mathbf{W}_s, \mathbf{W}_t} \sum_{c=1}^{C} \left\| \frac{1}{\hat{M}_s} \sum_{i=1}^{\hat{M}_s} \mathbf{W}_s^T \hat{\mathbf{x}}_i^{s(c)} - \frac{1}{\hat{M}_t^{(c)}} \sum_{j=1}^{\hat{M}_t^{(c)}} \mathbf{W}_t^T \hat{\mathbf{x}}_j^{t(c)} \right\|^2 \tag{7}$$

where $\hat{\mathbf{x}}_j^{t(c)}$ represents the j-th target sample within cluster c, and $\hat{M}_t^{(c)}$ denotes this cluster size. For simplicity, Eq. (7) can be reformulated by

$$\min_{\mathbf{W}_s, \mathbf{W}_t} \mathrm{Tr}\left([\mathbf{W}_s^T \ \mathbf{W}_t^T] \begin{bmatrix} \mathbf{M}_s & \mathbf{M}_{st} \\ \mathbf{M}_{ts} & \mathbf{M}_t \end{bmatrix} \begin{bmatrix} \mathbf{W}_s \\ \mathbf{W}_t \end{bmatrix} \right) \tag{8}$$

where $\begin{bmatrix} \mathbf{M}_s & \mathbf{M}_{st} \\ \mathbf{M}_{ts} & \mathbf{M}_t \end{bmatrix} \in \mathbb{R}^{2D\times 2D}$ is the MMD coefficient matrix. It is computed by

$$\begin{aligned} \mathbf{M}_s &= \hat{\mathbf{X}}_s \sum_{c=1}^{C} \mathbf{L}_s^{(c)} \hat{\mathbf{X}}_s^T, \quad \mathbf{M}_t = \hat{\mathbf{X}}_t \sum_{c=1}^{C} \mathbf{L}_t^{(c)} c \hat{\mathbf{X}}_t^T, \\ \mathbf{M}_{st} &= \hat{\mathbf{X}}_s \sum_{c=1}^{C} \mathbf{L}_{st}^{(c)} \hat{\mathbf{X}}_t^T, \quad \mathbf{M}_{ts} = \hat{\mathbf{X}}_t \sum_{c=1}^{C} \mathbf{L}_{ts}^{(c)} \hat{\mathbf{X}}_s^T. \end{aligned} \tag{9}$$

We assume that $\hat{\mathbf{X}}_s^{(c)}$ and $\hat{\mathbf{X}}_t^{(c)}$ denote the sets of source and target samples belonging to cluster c, respectively. The elements of $\mathbf{L}_s^{(c)}$ are computed by $(l_s^{(c)})_{ij} = 1/\hat{M}_s^2$ when $\hat{\mathbf{x}}_i^s, \hat{\mathbf{x}}_j^s \in \hat{\mathbf{X}}_s^{(c)}$. Similarly, elements of $\mathbf{L}_t^{(c)}$ are given by $(l_t^{(c)})_{ij} = 1/(\hat{M}_t^{(c)})^2$ when $\hat{\mathbf{x}}_i^t, \hat{\mathbf{x}}_j^t \in \hat{\mathbf{X}}_t^{(c)}$. For $\mathbf{L}_{st}^{(c)}$ and $\mathbf{L}_{ts}^{(c)}$, their elements are defined as $(l_{st}^{(c)})_{ij} = (l_{ts}^{(c)})_{ij} = -1/\hat{M}_s \hat{M}_t^{(c)}$ if $\hat{\mathbf{x}}_i^s \in \hat{\mathbf{X}}_s^{(c)}$ and $\hat{\mathbf{x}}_j^t \in \hat{\mathbf{X}}_t^{(c)}$. All other elements of these matrices are set to zero.

Overall Objective Function. By combining Eqs. (2), (5), (6) and (8), the overall objective function of the proposed DCSA is formulated by

$$\min_{\mathbf{W}_s, \mathbf{W}_t} \frac{\mathrm{Tr}(\begin{bmatrix} \mathbf{W}_s^T & \mathbf{W}_t^T \end{bmatrix} \begin{bmatrix} \mu\mathbf{D}_w + \gamma\mathbf{M}_s & \gamma\mathbf{M}_{st} \\ \gamma\mathbf{M}_{ts} & \alpha\mathbf{U} + \gamma\mathbf{M}_t \end{bmatrix} \begin{bmatrix} \mathbf{W}_s \\ \mathbf{W}_t \end{bmatrix})}{\mathrm{Tr}(\begin{bmatrix} \mathbf{W}_s^T & \mathbf{W}_t^T \end{bmatrix} \begin{bmatrix} \mu\mathbf{D}_b & 0 \\ 0 & \mathbf{V} \end{bmatrix} \begin{bmatrix} \mathbf{W}_s \\ \mathbf{W}_t \end{bmatrix})} \tag{10}$$

where $\mathbf{U} = \hat{\mathbf{X}}_t \mathbf{L}_t \hat{\mathbf{X}}_t^T$ and $\mathbf{V} = \hat{\mathbf{X}}_t \mathbf{H}_t \hat{\mathbf{X}}_t^T$ correspond to Eqs. (5) and (6), respectively. It should be noted that minimizing the numerator of Eq. (10) encourages small within-class variance in the source scene, preserves the geometric structure of the target scene, and reduces the conditional distribution discrepancy between scenes, while maximizing the denominator promotes large variance in the target scene and increased between-cluster variance in the source scene.

Algorithm 1: APGL

Input: Source data $\mathbf{X}_s$, source superpixel size M_s, projected dimension B, cluster size C, parameters λ, β

Perform superpixel segmentation on $\mathbf{X}_s$;
Initialize $\mathbf{P}$, $\mathbf{A}$, $\mathbf{Z}$, $\mathbf{F}$, $\mathbf{G}$;
repeat
    Update $\mathbf{Z}$ by solving Eq. (12);
    Update $\mathbf{F}$ by Eq. (14);
    Update $\mathbf{G}$ by Eq. (16);
    Update $\mathbf{P}$ by solving Eq. (18);
    Update $\mathbf{A}$ by Eq. (20);
until Convergence

Output: Label matrix of source data $\mathbf{Y}_s = \mathbf{ZG}$

Algorithm 2: DCSA

Input: Source data $\mathbf{X}_s$, clustering source labels $\mathbf{Y}_s$, target data $\mathbf{X}_t$, target superpixel size $\hat{N}_t$, subspace dimension R, per-cluster selected source sample size $\hat{M}_s$, parameters μ, α, γ

Select $\hat{M}_s$ samples from each cluster to form $\hat{\mathbf{X}}_s$;
Average pixels in each target superpixel to form $\hat{\mathbf{X}}_t$;
Train a classifier based on $\{\hat{\mathbf{X}}_s, \hat{\mathbf{Y}}_s, \hat{\mathbf{X}}_t\}$ to obtain $\hat{\mathbf{Y}}_t$;
Compute $\mathbf{D}_w$ and $\mathbf{D}_b$ by Eqs. (3) and (4);
Compute $\mathbf{L}_t$ and $\mathbf{H}_t$ by Eqs. (5) and (6);
Construct $\mathbf{M}_s$, $\mathbf{M}_t$, $\mathbf{M}_{st}$, and $\mathbf{M}_{ts}$ by solving Eq. (9);
Compute $\mathbf{W}_s$ and $\mathbf{W}_t$ by solving Eq. (23);

Output: Projection matrices $\mathbf{W}_s$ and $\mathbf{W}_t$

Source-to-Target Label Inference

Finally, a KNN algorithm (Cover and Hart 1967) is performed on the aligned source data $\mathbf{W}_s^T\hat{\mathbf{X}}_s$ with clustering labels $\hat{\mathbf{Y}}_s$ to infer the accurate labels $\mathbf{Y}_t$ for the aligned target data $\mathbf{W}_t^T\mathbf{X}_t$. This completes the source-to-target label inference based on the preceding steps.

Optimization

Optimization for APGL. To solve APGL in Eq. (1), we devise an alternating optimization algorithm to update each variable while keeping the remaining variables fixed.

Updating Z: When fixing $\mathbf{P}$, $\mathbf{A}$, $\mathbf{F}$ and $\mathbf{G}$, we have

$$\begin{aligned}\min_{\mathbf{Z}} \sum_{i=1}^{N_s}\sum_{j=1}^{M_s} \left\|\mathbf{P}^T\mathbf{x}_i^s - \mathbf{P}^T\mathbf{a}_j\right\|_2^2 z_{ij} + \beta\|\mathbf{Z}\|_F^2 \\ \text{s.t. } \mathbf{z}_i^T\mathbf{1} = 1, z_{ij} \geq 0, \mathbf{P}^T\mathbf{X}\mathbf{X}^\mathbf{T}\mathbf{P} = \mathbf{I}.\end{aligned} \tag{11}$$

Letting $d_{ij} = \left\|\mathbf{P}^T\mathbf{x}_i^s - \mathbf{P}^T\mathbf{a}_j\right\|_2^2$, Eq. (11) is rewritten by

$$\min_{\mathbf{z}_i^T\mathbf{1}=1, z_{ij}\geq 0} \frac{1}{2}\left\|\mathbf{z}_i + \frac{1}{2\beta}\mathbf{d}_i\right\|_2^2 \tag{12}$$

Eq. (12) can be solved with a closed-form solution (Nie et al. 2016).

Updating F: When fixing $\mathbf{P}$, $\mathbf{A}$, $\mathbf{G}$ and $\mathbf{Z}$, we have

$$\min_{\mathbf{F}} \|\mathbf{P}^T\mathbf{A} - \mathbf{F}\mathbf{G}^T\|_F^2, \text{ s.t. } g_{jk} \in \{0,1\}, \mathbf{g}_j^T\mathbf{1} = 1. \tag{13}$$

Taking the derivative of $\mathbf{F}$ and setting it to zero, $\mathbf{F}$ is updated by

$$\mathbf{F} = \mathbf{P}^T\mathbf{A}\mathbf{G}(\mathbf{G}^T\mathbf{G})^{-1} \tag{14}$$

Updating G: When fixing $\mathbf{P}$, $\mathbf{A}$ and $\mathbf{F}$, we have

$$\min_{\mathbf{G}} \|\mathbf{P}^T\mathbf{A} - \mathbf{F}\mathbf{G}^T\|_F^2, \text{ s.t } g_{jk} \in \{0,1\}, \mathbf{g}_j^T\mathbf{1} = 1. \tag{15}$$

This is a K-means-like problem that can be transformed into $\min\sum_{k=1}^{C}\sum_{a_j\in\mathbb{C}_k}\left\|\mathbf{P}^T\mathbf{a}_j - \mathbf{f}_k\right\|_2^2$, where $\mathbf{a}_j$ is assigned to the nearest cluster center $\mathbf{f}_k$. Thus, $\mathbf{G}$ can be updated element-wise as

$$g_{jk} = \begin{cases} 1, & k = \arg\min_l \left\|\mathbf{P}^T\mathbf{a}_j - \mathbf{f}_l\right\|_2^2, \\ 0, & \text{otherwise.}\end{cases} \tag{16}$$

Updating P: When fixing $\mathbf{A}$, $\mathbf{Z}$, $\mathbf{F}$ and $\mathbf{G}$, we have

$$\begin{aligned}\min_{\mathbf{P}} \sum_{i=1}^{N_s}\sum_{j=1}^{M_s} \left\|\mathbf{P}^T\mathbf{x}_i^s - \mathbf{P}^T\mathbf{a}_j\right\|_2^2 z_{ij} + \lambda\left\|\mathbf{P}^T\mathbf{A} - \mathbf{F}\mathbf{G}^T\right\|_F^2 \\ \text{s.t. } \mathbf{z}_i^T\mathbf{1} = 1, z_{ij} \geq 0, \mathbf{P}^T\mathbf{X}\mathbf{X}^T\mathbf{P} = \mathbf{I}, \\ g_{jk} \in \{0,1\}, \mathbf{g}_j^T\mathbf{1} = 1.\end{aligned} \tag{17}$$

Based on (Wang et al. 2021), Eq. (17) can be simplified to

$$\min_{\mathbf{P}^T\mathbf{X}\mathbf{X}^T\mathbf{P}=\mathbf{I}} \mathrm{Tr}\left[\mathbf{P}^T\left(\mathbf{Q}_1 + \mathbf{Q}_2\right)\mathbf{P}\right] \tag{18}$$

where $\mathbf{Q}_1 = \mathbf{X}\mathbf{D}_Z^{\text{row}}\mathbf{X}^T + \mathbf{A}\mathbf{D}_Z^{\text{col}}\mathbf{A}^T - 2\mathbf{X}\mathbf{Z}\mathbf{A}^T$, and $\mathbf{Q}_2 = \mathbf{A}\mathbf{D}_G^{\text{row}}\mathbf{A}^T + \hat{\mathbf{F}}\mathbf{D}_G^{\text{col}}\hat{\mathbf{F}}^T - 2\mathbf{A}\mathbf{G}\hat{\mathbf{F}}^T$. Additionally, $\mathbf{D}_Z^{\text{row}} = \mathrm{diag}(\sum_{j=1}^{M_s} z_{ij})$, $\mathbf{D}_Z^{\text{col}} = \mathrm{diag}(\sum_{i=1}^{N_s} z_{ij})$, $\mathbf{D}_G^{\text{row}} = \mathrm{diag}(\sum_{j=1}^{C} g_{kj})$, and $\mathbf{D}_G^{\text{col}} = \mathrm{diag}(\sum_{i=1}^{M_s} g_{kj})$. Thus, the optimal $\mathbf{P}$ is formed by the r eigenvectors of $\mathbf{Q}_1 + \mathbf{Q}_2$ corresponding to the r smallest eigenvalues.

Updating A: When fixing $\mathbf{Z}$, $\mathbf{F}$, $\mathbf{G}$ and $\mathbf{P}$, we have

$$\begin{aligned}\min_{\mathbf{A}} \sum_{i=1}^{N_s}\sum_{j=1}^{M_s} \left\|\mathbf{P}^T\mathbf{x}_i^s - \mathbf{P}^T\mathbf{a}_j\right\|_2^2 z_{ij} + \lambda\left\|\mathbf{P}^T\mathbf{A} - \mathbf{F}\mathbf{G}^T\right\|_F^2 \\ \text{s.t. } \mathbf{z}_i^T\mathbf{1} = 1, z_{ij} \geq 0, \mathbf{P}^T\mathbf{X}\mathbf{X}^T\mathbf{P} = \mathbf{I}, \\ g_{jk} \in \{0,1\}, \mathbf{g}_j^T\mathbf{1} = 1.\end{aligned} \tag{19}$$

Taking the derivative of $\mathbf{A}$ and setting it to zero, $\mathbf{A}$ is updated by

$$\mathbf{A} = \left(\mathbf{P}\mathbf{P}^T(\mathbf{X}\mathbf{Z} + \lambda\mathbf{F}\mathbf{G}^T)\right)\left(\mathbf{Z}^T\mathbf{Z} + \lambda\mathbf{I}\right)^{-1} \tag{20}$$

The entire optimization process for APGL is summarized in Algorithm 1.

Optimization for DCSA. To solve DCSA in Eq. (10), we reformulate it as a constrained optimization problem:

$$\begin{aligned}\min_{\mathbf{W}_s,\mathbf{W}_t} \mathrm{Tr}(\begin{bmatrix}\mathbf{W}_s^T & \mathbf{W}_t^T\end{bmatrix}\begin{bmatrix}\mu\mathbf{D}_w + \gamma\mathbf{M}_s & \gamma\mathbf{M}_{st} \\ \gamma\mathbf{M}_{ts} & \alpha\mathbf{U} + \gamma\mathbf{M}_t\end{bmatrix}\begin{bmatrix}\mathbf{W}_s \\ \mathbf{W}_t\end{bmatrix}) \\ \text{s.t. } \mathrm{Tr}(\begin{bmatrix}\mathbf{W}_s^T & \mathbf{W}_t^T\end{bmatrix}\begin{bmatrix}\mu\mathbf{D}_b & 0 \\ 0 & \mathbf{V}\end{bmatrix}\begin{bmatrix}\mathbf{W}_s \\ \mathbf{W}_t\end{bmatrix}) = 1.\end{aligned} \tag{21}$$

By introducing the Lagrange multiplier Φ, we have

$$\begin{aligned}\mathcal{L} = \mathrm{Tr}(\begin{bmatrix}\mathbf{W}_s^T & \mathbf{W}_t^T\end{bmatrix}\begin{bmatrix}\mu\mathbf{D}_w + \gamma\mathbf{M}_s & \gamma\mathbf{M}_{st} \\ \gamma\mathbf{M}_{ts} & \alpha\mathbf{U} + \gamma\mathbf{M}_t\end{bmatrix}\begin{bmatrix}\mathbf{W}_s \\ \mathbf{W}_t\end{bmatrix}) \\ + \mathrm{Tr}\Big((\begin{bmatrix}\mathbf{W}_s^T & \mathbf{W}_t^T\end{bmatrix}\begin{bmatrix}\mu\mathbf{D}_b & 0 \\ 0 & \mathbf{V}\end{bmatrix}\begin{bmatrix}\mathbf{W}_s \\ \mathbf{W}_t\end{bmatrix} - \mathbf{I})\Phi\Big)\end{aligned} \tag{22}$$

Tasks	Metrics	SGCNR	HESSC	S^{3}AGC	BGPC	SAPC	NCSC	SDST	AHSGC	SLCGC	Ours
Houston13→Houston18 (H13→H18)	ACC	0.4564	0.5396	0.6149	0.4210	0.5099	0.4598	0.4546	0.4771	0.4861	**0.8688**
	NMI	0.3063	0.3386	0.3222	0.2652	0.3294	0.1691	0.2016	0.1642	0.1670	**0.4645**
	Purity	0.6841	0.7041	0.6895	0.6601	0.6873	0.6239	0.6388	0.5056	0.5413	**0.8783**
Houston18→Houston13 (H18→H13)	ACC	0.5265	0.5632	0.5497	0.4783	0.5798	0.5834	0.4079	0.5272	0.5217	**0.8680**
	NMI	0.5260	0.5616	0.5988	0.4834	0.6207	0.4753	0.3394	0.3941	0.3667	**0.6456**
	Purity	0.5320	0.5731	0.5601	0.4794	0.5798	0.5834	0.4501	0.6213	0.5889	**0.8182**
PaviaU→PaviaC (PU→PC)	ACC	0.4492	0.6748	0.5492	0.6211	0.6411	0.6522	0.5091	0.5422	0.4848	**0.8341**
	NMI	0.2414	0.6162	0.5435	0.6605	0.5713	0.5279	0.3791	0.3607	0.2670	**0.6789**
	Purity	0.4565	0.6748	0.5600	0.6531	0.6878	0.6522	0.5374	0.5522	0.4968	**0.8139**
PaviaC→PaviaU (PC→PU)	ACC	0.5323	0.6141	0.6267	0.5744	0.5066	0.6285	0.4401	0.4690	0.4379	**0.7286**
	NMI	0.4509	0.4677	0.5286	0.5147	0.2692	0.3442	0.2496	0.1908	0.2426	**0.6314**
	Purity	0.6847	0.6555	0.6668	0.6910	0.5376	0.6511	0.5721	0.5522	0.5390	**0.8701**
Dioni→Loukia (Di→Lo)	ACC	0.5380	0.4692	0.6315	0.6264	0.5667	0.5369	0.3356	0.5365	0.4605	**0.8379**
	NMI	0.5305	0.4475	0.5650	0.5279	0.5217	0.4357	0.2684	0.4129	0.2994	**0.5777**
	Purity	0.6422	0.5812	0.6633	0.6350	0.6048	0.5514	0.4816	0.6603	0.5421	**0.8423**
Loukia→Dioni (Lo→Di)	ACC	0.5281	0.4623	0.4489	0.5147	0.4769	0.3738	0.3837	0.3982	0.3505	**0.7723**
	NMI	0.4991	0.4833	0.4328	0.4552	0.4422	0.2944	0.2293	0.3093	0.2144	**0.4992**
	Purity	0.6181	0.6414	0.5993	0.6376	0.5808	0.4478	0.4445	0.4605	0.3794	**0.7728**

Table 1: Performance of different methods for cross-scene clustering tasks. Best results are in **bold**, second-best are underlined.

By setting the derivative $\partial\mathcal{L}/\partial\left[\mathbf{W}_s^T\ \mathbf{W}_t^T\right]=0$, we have

$$\begin{bmatrix}\mu\mathbf{D}_w+\gamma\mathbf{M}_s & \gamma\mathbf{M}_{st}\\ \gamma\mathbf{M}_{ts} & \alpha\mathbf{U}+\gamma\mathbf{M}_t\end{bmatrix}\begin{bmatrix}\mathbf{W}_s\\ \mathbf{W}_t\end{bmatrix} = \Phi\begin{bmatrix}\mu\mathbf{D}_b & 0\\ 0 & \mathbf{V}\end{bmatrix}\begin{bmatrix}\mathbf{W}_s\\ \mathbf{W}_t\end{bmatrix}. \tag{23}$$

where $\Phi=\mathrm{diag}(\Phi_1,...,\Phi_k)$ contains the top k eigenvalues, and $[\mathbf{W}_s\ \mathbf{W}_t]$ is formed by the corresponding eigenvectors. It can be solved by generalized eigenvalue decomposition.

The entire optimization process for DCSA is summarized in Algorithm 2.

Time Complexity

The time complexity of ADSAC is dominated by the optimization of APGL and DCSA.

For APGL, updating each variable involves: $\mathcal{O}(BN_sM_s)$ for $\mathbf{Z}$, $\mathcal{O}(BDM_s)$ for $\mathbf{F}$, $\mathcal{O}(BDM_s)$ for $\mathbf{G}$, $\mathcal{O}(D^2N_s+DM_sN_s)$ for $\mathbf{P}$, and $\mathcal{O}(DN_sM_s+N_sM_s^2)$ for $\mathbf{A}$. With $C<B<D\approx M_s$ and $M_s\ll N_s$, the overall complexity is $\mathcal{O}(N_sDM_sT)$, nearly linear in N_s.

For DCSA, constructing matrices and performing eigenvalue decomposition takes: $\mathcal{O}(\hat{N}_s^2D^2)$ for $\mathbf{D}_w$ and $\mathbf{D}_b$, $\mathcal{O}(D\hat{N}_t^2+D^2\hat{N}_t)$ for $\mathbf{V}$ and $\mathbf{U}$, $\mathcal{O}(CD(\hat{N}_s^2+\hat{N}_t^2+\hat{N}_s\hat{N}_t))$ for the M matrices, and $\mathcal{O}(RD^2)$ for eigenvalue decomposition. The total complexity is $\mathcal{O}(\hat{N}_s^2D^2+CD\hat{N}_t^2)$.

Experiments

Experimental Setup

Datasets. There are three datasets adopted for experiments, including Houston, Pavia and HyRANK. The Houston dataset includes the Houston2013 (349×1905 pixels) and Houston2018 (349×1905 pixels) scenes, both containing seven land-cover categories and 48 spectral bands. The Pavia dataset consists of the Pavia Center (1096×715 pixels) and Pavia University (610×340 pixels) scenes, both with seven land covers and 102 spectral bands. The HyRANK dataset comprises the Dioni (250×1376 pixels) and Loukia (249×945 pixels) scenes, with 12 land-cover categories and 176 spectral bands. We evaluate our ADSAC on six cross-scene clustering tasks based on these datasets.

Compared methods. Since no existing cross-scene clustering methods are available, we compare our ADSAC with nine state-of-the-art general HSI clustering methods. These include five shallow methods: SGCNR (Wang et al. 2019), HESSC (Rafiezadeh Shahi et al. 2020), S^{3}AGC (Chen et al. 2023), BGPC (Zhang et al. 2024a), and SAPC (Jiang et al. 2025), along with four deep methods: NCSC (Cai et al. 2022), SDST (Luo et al. 2024), AHSGC (Ding et al. 2025a) and SLCGC (Ding et al. 2025b). Notably, clustering is first performed on the source scene, and the resulting labels are transferred to the target scene, as previously described.

Implementation Details. The optimal parameters for ADSAC is determined by a grid-search strategy. Specifically, B is decreased in steps of $D/6$, while R is selected from $[D-d, D-2d, ..., D-4d]$. Additionally, λ and α are varied from $[10^{-3}, 10^{-2}, 10^{-1}, 1, 10]$, μ is tuned within $[10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, 1, 10]$, and γ is selected from $[0.02:0.01:0.07]$, Moreover, M_s is chosen from $[250:25:350]$, $\hat{N}_t$ is selected from $[14000:500:16000]$. For other methods, we use their provided source code and optimal settings. To reduce randomness, all methods are run ten times, and average results are reported. Clustering performance is assessed using ACC, NMI, and purity. ADSAC and other shallow methods are implemented in MATLAB 2022b, while deep methods use Python 3.10. All experiments are conducted on a machine with a 3.00 GHz CPU and 64 GB RAM.

Results and Analyses

Clustering Performance Comparison. Table 1 shows the clustering performance of various methods in cross-scene tasks. ADSAC consistently outperforms all competitors, demonstrating strong generalization and robustness to distribution shifts. In Houston13→Houston18 and Houston18→Houston13, ADSAC achieves the highest accuracy. In more challenging tasks, it often exceeds the

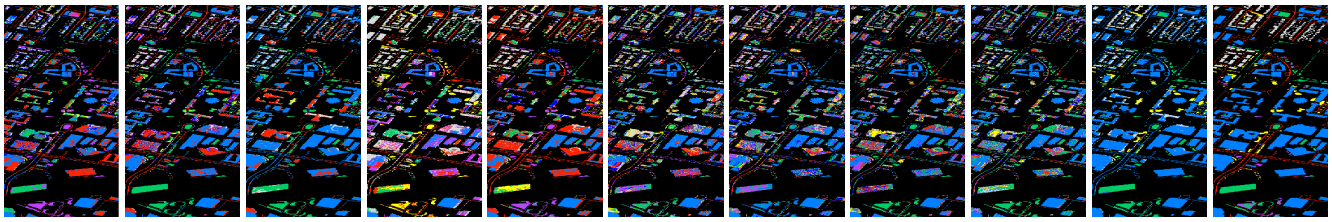

Figure 3: Clustering maps of Houston2018 via knowledge transfer from Houston2013 with diverse methods. GT: Ground truth.

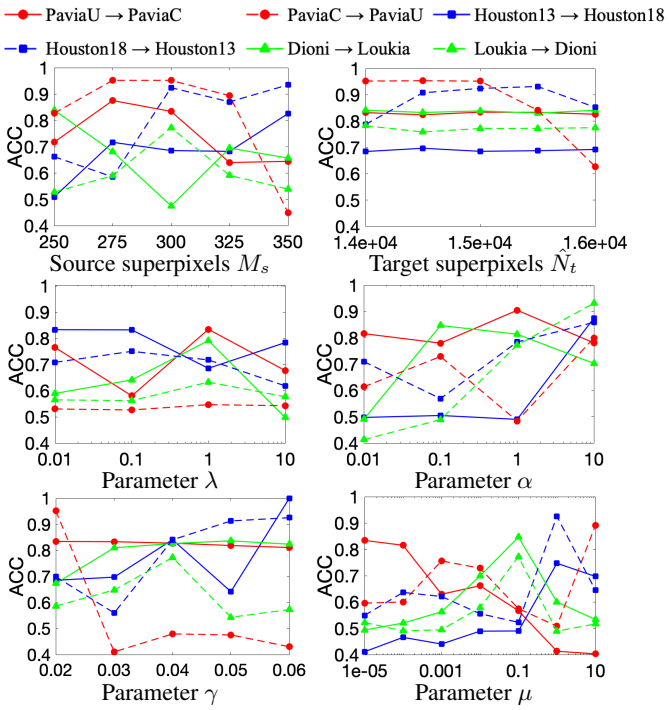

Figure 4: Parameter Sensitivity study of our ADSAC.

second-best method by a large margin, such as improving ACC by 16% and Purity by 14% over HESSC in PaviaU→PaviaC. It also outperforms S^{3}AGC by over 20% in Dioni→Loukia. These results highlight ADSAC's ability to effectively handle distribution shifts and align subspaces, making it a more flexible and reliable choice for cross-scene HSI clustering. As shown in Fig. 3, ADSAC's clustering maps most closely match the ground truth, further confirming its superior performance.

Parameter Sensitivity Analysis. Figure 4 illustrates the performance of ADSAC under different parameter settings. For the numbers of source and target superpixels (anchors) and the parameter λ, ACC varies only mildly within the tested ranges. For parameter α, ADSAC achieves the optimal ACC when $\alpha = 1$ in most cases. For parameter γ, ADSAC is less sensitive in PaviaU→PaviaC and Dioni→Loukia. For parameter μ, ADSAC shows fluctuant results in most cases. Fortunately, satisfactory results can be achieved with minimal parameter tuning.

Ablation Study. To assess ADSAC, we conduct an ablation study using three degraded variants. In these variants, APGL is replaced with k-means, and DCSA is removed to disable subspace alignment. As shown in Table 2, variant (I) discards both modules and yields the worst performance. Reintroducing either APGL or DCSA in variants (II) and (III) leads to clear accuracy gains. The full ADSAC model further outperforms all variants, demonstrating the advantage of jointly leveraging anchor-based source clustering and subspace alignment. We also replace APGL with four strong clustering baselines and observe consistently lower results, as shown in Table 3. This indicates that APGL offers more accurate source clustering for cross-scene transfer.

Settings		(I)	(II)	(III)	Ours
Modules	APGL	✗	✓	✗	✓
	DCSA	✗	✗	✓	✓
Tasks	H13→H18	0.3999	0.4638	0.5335	**0.8688**
	H18→H13	0.4272	0.5506	0.5725	**0.8418**
	PU→PC	0.5633	0.7008	0.6430	**0.8341**
	PC→PU	0.5233	0.6004	0.6211	**0.7286**
	Di→Lo	0.5236	0.6859	0.6250	**0.8379**
	Lo→Di	0.5163	0.6052	0.6972	**0.7723**

Table 2: Ablation study of our ADSAC (in ACC).

Tasks	S^{3}AGC	SAPC	SDST	SLCGC	Ours
H13→H18	0.6687	0.6657	0.5452	0.7645	**0.8688**
H18→H13	0.8497	0.7758	0.7060	0.7677	**0.8680**
PU→PC	0.7162	0.7086	0.6797	0.7669	**0.8341**
PC→PU	0.6491	0.6427	0.6669	0.5938	**0.7286**
Di→Lo	0.6355	0.7930	0.6836	0.6442	**0.8379**
Lo→Di	0.6113	0.7521	0.7182	0.6712	**0.7723**

Table 3: Effect of APGL (in ACC).

Conclusion

In this paper, we first proposed the ADSAC for cross-scene HSI clustering, comprising three key learning steps. First, the developed APGL model leverages anchor-based structural information to obtain accurate clustering labels for the source scene. Next, DCSA model learns discriminative feature representations for both scenes within aligned subspaces by reducing distribution discrepancies. Finally, a KNN classifier infers target labels by training on clustering-labeled source data and target data within aligned subspaces. To solve the formulated models, customized optimization algorithms are designed as solutions. Extensive experiments demonstrate the superiority of the proposed ADSAC.

Acknowledgments

This work was funded by the China National Key R&D Program (Grant 2023YFF0807000).

References

Cai, Y.; Zhang, Z.; Ghamisi, P.; Ding, Y.; Liu, X.; Cai, Z.; and Gloaguen, R. 2022. Superpixel contracted neighborhood contrastive subspace clustering network for hyperspectral images. *IEEE Transactions on Geoscience and Remote Sensing*, 60: 1–13.

Chen, J.; Hou, D.; He, C.; Liu, Y.; Guo, Y.; and Yang, B. 2024. Change detection with cross-domain remote sensing images: A systematic review. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 17: 11563–11582.

Chen, X.; Zhang, Y.; Feng, X.; Jiang, X.; and Cai, Z. 2023. Spectral-spatial superpixel anchor graph-based clustering for hyperspectral imagery. *IEEE Geoscience and Remote Sensing Letters*, 20: 1–5.

Cover, T.; and Hart, P. 1967. Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, 13(1): 21–27.

Ding, K.; Lu, T.; Fu, W.; and Fang, L. 2024. Cross-scene hyperspectral image classification with consistency-aware customized learning. *IEEE Transactions on Circuits and Systems for Video Technology*.

Ding, Y.; Kang, W.; Yang, A.; Zhang, Z.; Zhao, J.; Feng, J.; Hong, D.; and Zheng, Q. 2025a. Adaptive Homophily Clustering: Structure Homophily Graph Learning with Adaptive Filter for Hyperspectral Image. *IEEE Transactions on Geoscience and Remote Sensing*.

Ding, Y.; Zhang, Z.; Yang, A.; Cai, Y.; Xiao, X.; Hong, D.; and Yuan, J. 2025b. SLCGC: A lightweight Self-supervised Low-pass Contrastive Graph Clustering Network for Hyperspectral Images. *IEEE Transactions on Multimedia*.

Fisher, R. A. 1936. The use of multiple measurements in taxonomic problems. *Annals of Eugenics*, 7(2): 179–188.

Ghamisi, P.; Yokoya, N.; Li, J.; Liao, W.; Liu, S.; Plaza, J.; Rasti, B.; and Plaza, A. 2017. Advances in hyperspectral image and signal processing: A comprehensive overview of the state of the art. *IEEE Geoscience and Remote Sensing Magazine*, 5(4): 37–78.

Gretton, A.; Borgwardt, K. M.; Rasch, M. J.; Schölkopf, B.; and Smola, A. 2012. A kernel two-sample test. *The Journal of Machine Learning Research*, 13(1): 723–773.

Jiang, G.; Zhang, Y.; Wang, X.; Jiang, X.; and Zhang, L. 2025. Structured Anchor Learning for Large-Scale Hyperspectral Image Projected Clustering. *IEEE Transactions on Circuits and Systems for Video Technology*, 35(3): 2328–2340.

Kouw, W. M.; and Loog, M. 2019. A review of domain adaptation without target labels. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(3): 766–785.

Li, J.; Zhang, Z.; Song, R.; Li, Y.; and Du, Q. 2024a. SCFormer: Spectral coordinate transformer for cross-domain few-shot hyperspectral image classification. *IEEE Transactions on Image Processing*, 33: 840–855.

Li, Q.; Wen, Y.; Zheng, J.; Zhang, Y.; and Fu, H. 2024b. Hyunida: Breaking label set constraints for universal domain adaptation in cross-scene hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 62: 5518415.

Li, S.; Song, W.; Fang, L.; Chen, Y.; Ghamisi, P.; and Benediktsson, J. A. 2019. Deep learning for hyperspectral image classification: An overview. *IEEE Transactions on Geoscience and Remote Sensing*, 57(9): 6690–6709.

Liu, M.-Y.; Tuzel, O.; Ramalingam, S.; and Chellappa, R. 2011. Entropy rate superpixel segmentation. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2097–2104. IEEE.

Luo, F.; Liu, Y.; Duan, Y.; Guo, T.; Zhang, L.; and Du, B. 2024. SDST: Self-supervised double-structure transformer for hyperspectral images clustering. *IEEE Transactions on Geoscience and Remote Sensing*.

Ma, L.; Liu, Y.; Zhang, X.; Ye, Y.; Yin, G.; and Johnson, B. A. 2019. Deep learning in remote sensing applications: A meta-analysis and review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 152: 166–177.

Maćkiewicz, A.; and Ratajczak, W. 1993. Principal components analysis (PCA). *Computers & Geosciences*, 19(3): 303–342.

Nie, F.; Wang, X.; Jordan, M.; and Huang, H. 2016. The constrained laplacian rank algorithm for graph-based clustering. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 30.

Plaza, A.; Benediktsson, J. A.; Boardman, J. W.; Brazile, J.; Bruzzone, L.; Camps-Valls, G.; Chanussot, J.; Fauvel, M.; Gamba, P.; Gualtieri, A.; et al. 2009. Recent advances in techniques for hyperspectral image processing. *Remote Sensing of Environment*, 113: S110–S122.

Rafiezadeh Shahi, K.; Khodadadzadeh, M.; Tusa, L.; Ghamisi, P.; Tolosana-Delgado, R.; and Gloaguen, R. 2020. Hierarchical sparse subspace clustering (HESSC): An automatic approach for hyperspectral image analysis. *Remote Sensing*, 12(15): 2421.

Sohn, K.; Liu, S.; Zhong, G.; Yu, X.; Yang, M.-H.; and Chandraker, M. 2017. Unsupervised domain adaptation for face recognition in unlabeled videos. In *Proceedings of the IEEE International Conference on Computer Vision*, 3210–3218.

Tang, X.; Li, C.; and Peng, Y. 2022. Unsupervised joint adversarial domain adaptation for cross-scene hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 60: 1–15.

Tenenbaum, J. B.; Silva, V. d.; and Langford, J. C. 2000. A global geometric framework for nonlinear dimensionality reduction. *Science*, 290(5500): 2319–2323.

Wang, J.; Wang, L.; Nie, F.; and Li, X. 2021. Fast unsupervised projection for large-scale data. *IEEE Transactions on Neural Networks and Learning Systems*, 33(8): 3634–3644.

Wang, R.; Nie, F.; Wang, Z.; He, F.; and Li, X. 2019. Scalable graph-based clustering with nonnegative relaxation for large hyperspectral image. *IEEE Transactions on Geoscience and Remote Sensing*, 57(10): 7352–7364.

Xie, Z.; Duan, P.; Kang, X.; Liu, W.; and Li, S. 2024. Class-wise Prototype Guided Alignment Network for Cross-Scene Hyperspectral Image Classification. *IEEE Transactions on Geoscience and Remote Sensing*.

Xu, M.; Wu, M.; Chen, K.; Zhang, C.; and Guo, J. 2022. The eyes of the gods: A survey of unsupervised domain adaptation methods based on remote sensing data. *Remote Sensing*, 14(17): 4380.

Ye, M.; Ling, J.; Huo, W.; Zhang, Z.; Xiong, F.; and Qian, Y. 2024. Discriminative Vision Transformer for Heterogeneous Cross-Domain Hyperspectral Image Classification. *IEEE Transactions on Geoscience and Remote Sensing*.

Zhang, J.; Zhang, Y.; and Zhou, Y. 2023. Quantum-inspired spectral-spatial pyramid network for hyperspectral image classification. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 9925–9934.

Zhang, Y.; Jiang, G.; Cai, Z.; and Zhou, Y. 2024a. Bipartite graph-based projected clustering with local region guidance for hyperspectral imagery. *IEEE Transactions on Multimedia*, 26: 9551–9563.

Zhang, Y.; Li, W.; Sun, W.; Tao, R.; and Du, Q. 2023. Single-source domain expansion network for cross-scene hyperspectral image classification. *IEEE Transactions on Image Processing*, 32: 1498–1512.

Zhang, Y.; Li, W.; Tao, R.; Peng, J.; Du, Q.; and Cai, Z. 2021. Cross-scene hyperspectral image classification with discriminative cooperative alignment. *IEEE Transactions on Geoscience and Remote Sensing*, 59(11): 9646–9660.

Zhang, Y.; Yan, S.; Zhang, L.; and Du, B. 2024b. Fast projected fuzzy clustering with anchor guidance for multimodal remote sensing imagery. *IEEE Transactions on Image Processing*, 33: 4640–4653.

Zhao, C.; Qin, B.; Feng, S.; Zhu, W.; Zhang, L.; and Ren, J. 2022. An unsupervised domain adaptation method towards multi-level features and decision boundaries for cross-scene hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 60: 1–16.