

Boosting Universal Domain Adaptation in Remote Sensing With Dual-Classifiers Consistency Discrimination and Cross-Domain Feature Mixup

Qingmei Li^{ID}, Yang Zhang^{ID}, Juepeng Zheng^{ID}, Member, IEEE, Yuxiang Zhang, Jianxi Huang^{ID}, Senior Member, IEEE, and Haohuan Fu^{ID}, Senior Member, IEEE

Abstract—In the field of remote sensing (RS) image classification, domain adaptation (DA) methods have been extensively utilized to overcome the challenges posed by data discrepancies between source and target domains that arise from varying imaging conditions, sensor differences, or geographical variations. Stemming from the existence of unseen classes in both the source and target domains, universal DA (UniDA) poses the greatest challenge that demands innovative solutions. Existing UniDA methods often overlook intra-domain variations within the target domain and face difficulties in distinguishing between similar known and unknown classes, which significantly hinder cross-domain transfer. To overcome these challenges, we propose a dual-classifier network tailored for cross-domain classification of RS images, named DCmix. DCmix introduces a dual-classifiers network that utilizes both closed-set and open-set classifiers to improve the accuracy of identifying unknown sample classes. To our knowledge, this is the first attempt to introduce dual

classifiers into the UniDA RS image classification task. We further enhance the feature generalization capability of the target domain based on sample neighborhood relations, resulting in a more adaptable and robust feature representation. A cross-domain feature mixup (FM) scheme is also designed based on the consistency discrimination of the dual classifiers, achieving smoother decision boundaries and simpler hidden layer representations. Extensive experiments conducted on four hyperspectral image datasets and three RGB datasets prove that the introduced approach attains state-of-the-art (SOTA) performance in RS image classification under the UniDA scenario.

Index Terms—Consistency discrimination, cross-domain image classification, feature mixup (FM), remote sensing (RS), universal domain adaptation (UniDA).

I. INTRODUCTION

REMOTE sensing image classification serves as a pivotal tool for fine-grained analysis of the Earth's surface with applications spanning land use mapping [1], agricultural management [2], [3], and disaster monitoring [4]. The advent of deep learning (DL) techniques has brought significant advances in remote sensing (RS) image classification [5], [6], such as deep belief networks (DBNs) [7], convolutional neural networks (CNNs) [8], [9], generative adversarial networks (GANs) [10], stacked auto-encoder (SAE) [11], and Transformer [12]. Traditional methods rely heavily on handcrafted feature extraction, which is both time-consuming and often limited in performance. In contrast, DL methods automatically extract multilayered feature representations, significantly enhancing classification accuracy. For instance, CNNs [8], [9] effectively capture local features in RS images through convolutional layers, enabling deep understanding of complex scenes. GANs [10] employ adversarial training between a generator and a discriminator to generate high-quality RS image samples, thereby alleviating the scarcity of labeled data and enhancing the generalization capabilities of models. Despite these significant advances, DL methods often need extensive labeled data for training, and their performance tends to degrade when applied to domains with different characteristics, such as changes in image acquisition conditions or geographic locations, posing challenges for cross-domain classification.

Received 20 December 2024; revised 12 April 2025; accepted 14 May 2025. Date of publication 20 May 2025; date of current version 2 June 2025. This work was supported in part by the National Natural Science Foundation of China under Grant T2125006 and Grant 42401415; in part by Shenzhen Science and Technology Program under Grant KCXFZ20240903093759004 and Grant KJZD20230923115106012; in part by the Fundamental Research Funds for the Central Universities, Sun Yat-sen University under Project 24xkjc002; and in part by Jiangsu Innovation Capacity Building Program under Project BM2022028. (*Corresponding author: Juepeng Zheng.*)

Qingmei Li is with Tsinghua Shenzhen International Graduate School, Tsinghua University, Shenzhen 518071, China (e-mail: qingmeili1997@gmail.com).

Yang Zhang is with the School of Artificial Intelligence, Sun Yat-sen University, Zhuhai 519082, China (e-mail: zhangy2583@mail2.sysu.edu.cn).

Juepeng Zheng is with the School of Artificial Intelligence, Sun Yat-sen University, Zhuhai 519082, China, and also with the National Supercomputing Center in Shenzhen, Shenzhen 518117, China (e-mail: zhengjp8@mail.sysu.edu.cn).

Yuxiang Zhang is with the School of Information and Electronics, Beijing Institute of Technology, Beijing 100811, China (e-mail: zyx829625@163.com).

Jianxi Huang is with the Faculty of Geosciences and Engineering, Southwest Jiaotong University, Chengdu 610032, China, also with the College of Land Science and Technology, China Agricultural University, Beijing 100107, China, and also with the Key Laboratory of Remote Sensing for Agri-Hazards, Ministry of Agriculture and Rural Affairs, Beijing 100125, China (e-mail: jxhuang@cau.edu.cn).

Haohuan Fu is with Tsinghua Shenzhen International Graduate School, Tsinghua University, Shenzhen 518071, China, also with the National Supercomputing Center in Shenzhen, Shenzhen 518117, China, and also with the Ministry of Education Key Laboratory for Earth System Modeling, Department of Earth System Science, Tsinghua University, Beijing 100084, China (e-mail: haohuan@tsinghua.edu.cn).

Digital Object Identifier 10.1109/TGRS.2025.3571747

To bridge the gap in the cross-domain scenario, domain adaptation (DA) has emerged as a solution, facilitating the effective transfer of knowledge acquired from a labeled domain to an unfamiliar domain. DA methods can be broadly classified into four categories: closed-set DA [13], [14], [15], partial DA [16], [17], [18], open-set DA [19], [20], and universal DA (UniDA) [21], [22], [23], [24]. Closed-set DA refers to scenarios where the source and target domains include identical classes. Shi et al. [14] proposed an unsupervised DA method for SAR target classification, aligning simulated and real domain distributions using gradient-weighted adversarial alignment and prototypical network-based class-level alignment, enhanced with contrastive learning and pseudo-label filtering. Partial DA assumes that the source domain contains categories that are not present in the target domain, while open-set DA considers that the target domain may include classes that are absent in the source domain. Ma et al. [18] introduced the partial domain adversarial neural network, which mitigates negative transfer by down-weighting outlier source samples and performing partial weighted alignment between target and source domains for crop yield prediction. Zheng et al. [20] investigated the open-set DA scenarios in the RS cross-scene classification by employing a multiadversarial open-set DA network that integrates attention-aware open-set backpropagation, auxiliary adversarial learning, and adaptive entropy suppression. UniDA poses the most challenging scenario, as it aims to handle the general case where both partial overlap and unknown classes may exist.

UniDA scenario is really common in real-world RS applications. For instance, when we conduct a forest inventory using RS images (tree species classification), exploring specific tree species in advance remains a time-consuming and laborious task. Within the field of RS, despite the various methods proposed for DA scenarios, UniDA for RS images is still not receiving much attention. Xu et al. [22] introduced the UniDA approach for RGB image classification, model adaptation (MA), and source data generation–MA (SDG-MA), which involves setting a manual threshold to filter out unknown instances, as shown in Fig. 1(a). Our previous work, HyUniDA, achieved remarkable transfer learning for hyperspectral images (HSIs) [24]; however, we noticed that the clustering accuracy in the target domain can be influenced by the choice of initial conditions. Most existing methods only focus on feature alignment and ignore the interaction learning between classifiers, which is crucial to improving the classification performance of unknown categories. As Fig. 1(b) and (c) describes, one-vs-all classifiers train a binary classifier for each category, learning to distinguish between positive and negative classes, thus simplifying the decision-making process and making it more robust to the presence of unknown classes. Saito and Saenko [25] proposed the OVNNet based on one-vs-all classifiers, using the inter-class distance among source classes to determine the threshold. However, an application of entropy minimization as a sole criterion for classification can result in the failure to properly identify unknown classes, potentially misleading the analysis by misdirecting attention toward incorrectly classified samples.

To overcome these challenges, we introduce a method that leverages one-vs-all (OVA) classifiers for UniDA in

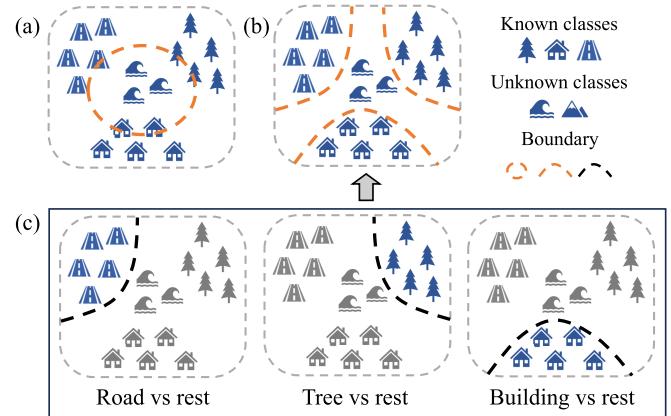


Fig. 1. Approaches of existing UniDA methods and one-vs-all classifiers in handling unknown instances. (a) Manual threshold definition for unknown instance rejection. (b) and (c) How OVA classifiers work by training a binary classifier for each category and learning to distinguish between positive and negative classes, respectively.

RS image classification, named DCmix, which combines a dual-classifiers consistency discrimination network with a cross-domain feature mixup (FM) strategy. The diverse environmental conditions, sensor types, or imaging angles cause significant discrepancies across different domains (i.e., inter-domain variations), even within the same class in a given domain, due to wide coverage (i.e., intra-domain variations). We identify reliable neighbors for each sample, maximize their pairwise similarity to minimize intra-domain variations, and measure neighbor confidence to acquire invariant features that effectively generalize to the target domain. However, relying solely on feature representation learning is insufficient to fully mitigate the effect of category shift, especially in the presence of unknown class samples. Therefore, we propose a cross-domain FM approach, which explicitly simulates unknown samples by utilizing arbitrary intermediate states of mixups between domains, smoothing the transfer process, and enhancing the identification of unknown classes. Nevertheless, FM may lead to misidentification of known class instances. To address the issue, we design the dual-classifiers consistency discrimination that establishes consistency between closed-set and open-set classifiers, aiming to optimize known category samples that are wrongly identified as unknown while maintaining the capability to detect unknown samples. Through combining these techniques, our proposed DCmix is able to more effectively distinguish similar items from both known and unknown categories in the target domain, thus enhancing overall classification capability. Our experiments on four HSI datasets and three RGB datasets highlight its excellence, achieving state-of-the-art (SOTA) results.

To sum up, this work provides the following key contributions to the RS field.

- 1) We propose a dual-classifier network tailored for cross-domain RS images classification, leveraging both open-set classifiers and closed-set classifiers to enhance the accuracy of identifying unknown sample classes. To our knowledge, this is the first work to introduce the dual-classifiers to the DA task of RS image classification.

- 2) We enhance the feature generalization ability of the target domain based on the neighborhood relation of samples, which involves the processes of reliable neighbor search, similarity maximization, and confidence measurement, ultimately leading to a more adaptable and robust feature representation within the target domain.
- 3) We develop a cross-domain FM scheme based on the constraints of dual-classifiers consistency discrimination, and achieve smoother decision boundaries and simpler hidden layer representations by explicitly simulating any intermediate state of the unknown class.

The subsequent sections are structured as follows. In Section II, we briefly review the related works in the DA field. Section III provides an in-depth explanation of the proposed method, followed by a description of our datasets in Section IV. The performance evaluation, along with extensive experiments and analyses, is presented in Section V. Lastly, we conclude the article in Section VI.

II. RELATED WORK

A. Universal DA

UniDA describes a situation where both the source and target domains share some common labels, but each domain also contains private labels not found in the other [21], [26], [27]. The objective of UniDA is to correctly classify target samples either into one of the known labels or an unknown label without any prior knowledge. You et al. [21] proposed the idea of UniDA and designed UAN to discover the shared and private label sets by exploiting both domain similarity and prediction uncertainty for individual samples. CMU and H-score were proposed by Fu et al. [28] to improve accuracy in open class detection. All the approaches mentioned above manually set and adjust a threshold to detect common and unknown samples, which limits their extension to more practical scenarios. To address this issue, Chang et al. [29] proposed UniOT to find common classes without a hand-tuned threshold and used global statistical information of the assignment matrix to distinguish common and private classes internally. In RS community, Xu et al. [22] introduced a UniDA approach for RGB image classification, including MA DA SDG-MA. C³DA proposed by Guo et al. [23] has an ensemble criterion for “unknown” classes fusing confidence, consistency, and certainty of samples to achieve higher performance under the UniDA scenario. Our previous work HyUniDA achieved remarkable transfer performance for HSIs based on the shared semantic pairing and domain similarity score [24]. In distinction to prior advancements, our method proposes a dual-classifier network for cross-domain RS image classification, including feature alignment, mixup, and consistency discrimination.

B. Mixup Approaches

Data augmentation techniques, particularly mixup, have gained significant attention in the field of unsupervised DA (UDA). Initially proposed by Zhang et al. [30], mixup generates synthetic samples by linearly interpolating between pairs of training examples, which enhances model generalization. Prior studies [31], [32] utilized classifier predictions

as pseudo-labels for target instances, and incorporated vanilla mixup within the target domain to enhance prediction stability and robustness. Diverging from the aforementioned technique, Tranheden et al. [33] extended this idea for semantic segmentation with DACS by replacing vanilla mixup with ClassMix [34], which merges pixels from different classes and positions. Manifold mixup [35] enhances model performance by learning smoother decision boundaries, capturing high-level information through deep layer interpolation, and flattening class representations.

The consistency constraint indeed plays a critical role in DA. Different consistency training methods vary in how to generate data perturbation and how the consistency loss is composed. Methods such as AutoAugment [36], population-based augmentation (PBA) [37], and RandAugment [38] use policy search to generate optimal data augmentation strategies, thereby enhancing model consistency. By using consistent prediction as a constraint, French et al. [39] integrated a mean teacher model to achieve DA. Notably, Xiao et al. [40] employed data augmentation techniques with image rotation to establish a connection between self-supervised learning and consistency learning.

Despite these advancements, these methods are not directly applicable to UniDA due to category-shift issues, where simulating unknown-class samples is crucial. To address this, we introduce a cross-domain FM scheme, where mixup features exhibit a low probability of belonging to known categories, enhancing the identification of unknown classes. The consistency constraint maximizes the mutual information between the transformed data and the main task labels.

C. OVA Classifiers

The OVA classifiers involve training a separate binary classifier for each class, where each classifier distinguishes one class from all other classes [41]. This strategy is primarily employed to extend binary classifiers for multiclass classification problems, like support vector machines, logistic regression, etc. [41], [42]. OVA approaches have been found to uncover more pertinent hidden representations for unidentified samples compared to the widely used Softmax function, thus improving the accuracy and interpretability of DL models [43]. Padhy et al. [44] revisited OVA classifiers to enhance prediction uncertainty and out-of-distribution (OOD) detection, showcasing that OVA can better calibrate predictions compared to the standard softmax approach. Saito and Saenko [25] proposed OVANet to avoid manually setting thresholds for rejecting unknown samples in UniDA by letting the open-set classifiers focus on hard negative samples. Based on the framework of OVANet [25], we further explore the correlation between closed-set and open-set classifiers to optimize UniDA’s performance.

III. METHODOLOGY

In this section, we present a thorough explication of the designed DCmix model, including three modules: feature alignment based on neighborhood relation, cross-domain FM, and dual-classifiers consistency discrimination, as well as the overall objective of the model. The feature alignment module

aims to reduce intra-domain variations by identifying reliable neighbors for each sample and maximizing their pairwise similarity. Then, the cross-domain FM module explicitly simulates unknown categories by generating intermediate feature representations between source and target domains, enhancing the identification of unknown classes. Finally, by enforcing alignment between closed-set and open-set classifiers, the dual-classifiers consistency discrimination module improves the classification accuracy of known samples while preserving the model's ability to identify unknown samples. Fig. 2 presents an outline of the proposed DCmix.

A. Problem Setting

In the UniDA scenario, we face the challenge of classifying unknown target domain $\mathcal{D}^t = \{\mathbf{x}_i^t\}_{i=1}^{n_t}$ using labeled source domain $\mathcal{D}^s = \{\mathbf{x}_i^s, \mathbf{y}_i^s\}_{i=1}^{n_s}$. The source data consists of n_s labeled samples, while the target data has n_t unlabeled samples. \mathcal{C}_s and \mathcal{C}_t represent the label sets for source and target domains, respectively. $\mathcal{C} = \mathcal{C}_s \cup \mathcal{C}_t$ signifies the label set encompassing all classes from both domains. We aim to categorize each target instance into a class from the source class set \mathcal{C}_s or the “unknown” class set $\bar{\mathcal{C}}_t = \mathcal{C}_t \setminus \mathcal{C}$ that represents classes exclusive to the target domain.

B. Feature Alignment

If feature embeddings are properly trained on both source and target domains, the features of target samples located within adjacent regions are likely to belong to the identical category. The local regions that preserve this homogeneity are defined as invariant spaces in this study. By exploiting the invariance of neighboring samples, we mitigate cross-domain variations and learn the discriminative features to more effectively distinguish classes within the target domain.

The features of all target samples are stored in a memory bank $\mathcal{M} = \mathbb{R}^{H \times W \times B \times n_t}$. The entry $\mathbf{e}_i \in \mathbb{R}^{H \times W \times B}$ contains the feature associated with target sample x_i^t , which undergoes a dynamic updating process after each mini-batch cycle and unified via L2 normalization to maintain stability. The L2 normalization ensures that all feature vectors have the same scale, reducing the influence of numerical imbalances during training and enhancing the stability of the alignment process

$$\mathbf{e}_i \leftarrow \frac{\mathcal{F}(x_i^t)}{\|\mathcal{F}(x_i^t)\|_2} \quad (1)$$

where $\mathcal{F}(x_i^t)$ is the feature extracted by the classifier. The probability that x_j^t and x_i^t are classified into the same category can be calculated as

$$P_{ij} = \frac{\exp(\delta \cdot \mathbf{e}_i \cdot \mathbf{e}_j)}{\sum_{h=1}^{n_t} \exp(\delta \cdot \mathbf{e}_i \cdot \mathbf{e}_h)} \quad (2)$$

where \mathbf{e}_h represents the feature vector of the h th sample (x_h^t) in the target domain, δ acts as a scale parameter that regulates the shape and characteristics of probability distributions, which is set to 10 [45]. We mitigate intra-domain variations through the maximization of probabilities associated with each input instance and its adjacent samples.

Nevertheless, the efficacy of this procedure relies on the accuracy of neighbor search, which is essential for preserving

the neighborhood invariance of embedding space within the target domain. It is common to select the K -nearest neighbors as neighborhood [46], [47], [48], but it is limited by the quantity of samples per class, and determining an apt K -value is a costly process for every dataset or class.

We propose an adaptive neighborhood search strategy. As shown in Fig. 3(a) and (b), the neighbor of x_i^t can be defined as

$$\Omega_i = \{\mathbf{e}_i \cdot \mathbf{e}_k > \psi \cdot \max(\mathbf{e}_i \cdot \mathbf{e}_{n_0}), k \neq i\} \quad (3)$$

where n_0 is the nearest neighbor of x_i^t . ψ is the neighbor similarity, which is set to 0.875. Our neighborhood selection framework exhibits stability and robustness against imbalances in data size, as detailed proof is given in the Appendix.

Utilizing the Jaccard distance [49], we gauge the weight to which the j th instance in memory can be considered a neighboring sample of the x_i^t , as shown in Fig. 3(c)

$$\mathcal{W}_{ij} = \frac{|\Omega_i \cap \Omega_j|}{|\Omega_i \cup \Omega_j|} \quad (4)$$

where $|\Omega_i \cap \Omega_j|$ denotes the number of common neighbors between sample i and j , and $|\Omega_i \cup \Omega_j|$ is the size of the union of their neighbor sets. Jaccard distance measures the similarity of neighborhood structures between samples. A smaller distance indicates that samples i and j share highly similar neighborhoods, implying greater intra-class consistency.

The feature alignment based on neighbor confidence is explicated as follows:

$$\mathcal{L}_{FA}(x_i^t) = -\frac{1}{|\Omega_i|} \sum_{j \in \Omega_i} \mathcal{W}_{ij} \log P_{ij}. \quad (5)$$

The process of minimizing the loss function $\mathcal{L}_{FA}(x_i^t)$ imposes a similarity constraint on the feature representations of reliable neighbors with high confidence, which serves to decrease variations within the domain.

C. Feature Mixup

The spectral information of RS image varies with the season and weather conditions of data acquisition, resulting in a shift between the source and target domains. Despite feature alignment based on neighbor confidence can enhance the discriminability of feature representations, the challenge of class-shift persists. In OVANet [25], the hard-negative classifier sampling primarily focuses on inter-class distances within the source domain, the open-set entropy minimization misclassifies several target samples from unlabeled classes as classified into one of the predefined categories. To overcome this limitation and mitigate distribution shift, we introduce a cross-domain FM approach inspired by manifold mixup [35], to smoothly simulate unknown-class samples across domains.

Given the samples x_i^s and x_j^t with the corresponding feature representations at layer given by $\mathcal{F}_\theta^l(x_i^s)$ and $\mathcal{F}_\theta^l(x_j^t)$, the mixup feature can be formulated as

$$\begin{aligned} \text{Mix}(x_i^s, x_j^t) &= \text{Mix}(\mathcal{F}_\theta^l(x_i^s), \mathcal{F}_\theta^l(x_j^t)) \\ &= \lambda \cdot \mathcal{F}_\theta^l(x_i^s) + (1 - \lambda) \cdot \mathcal{F}_\theta^l(x_j^t) \end{aligned} \quad (6)$$

where $\lambda \sim \text{Beta}(\alpha, \alpha)$, and α is fixed at 2.0 for the Beta distribution. Considering the interpolation parameter $\lambda \in [0, 1]$,

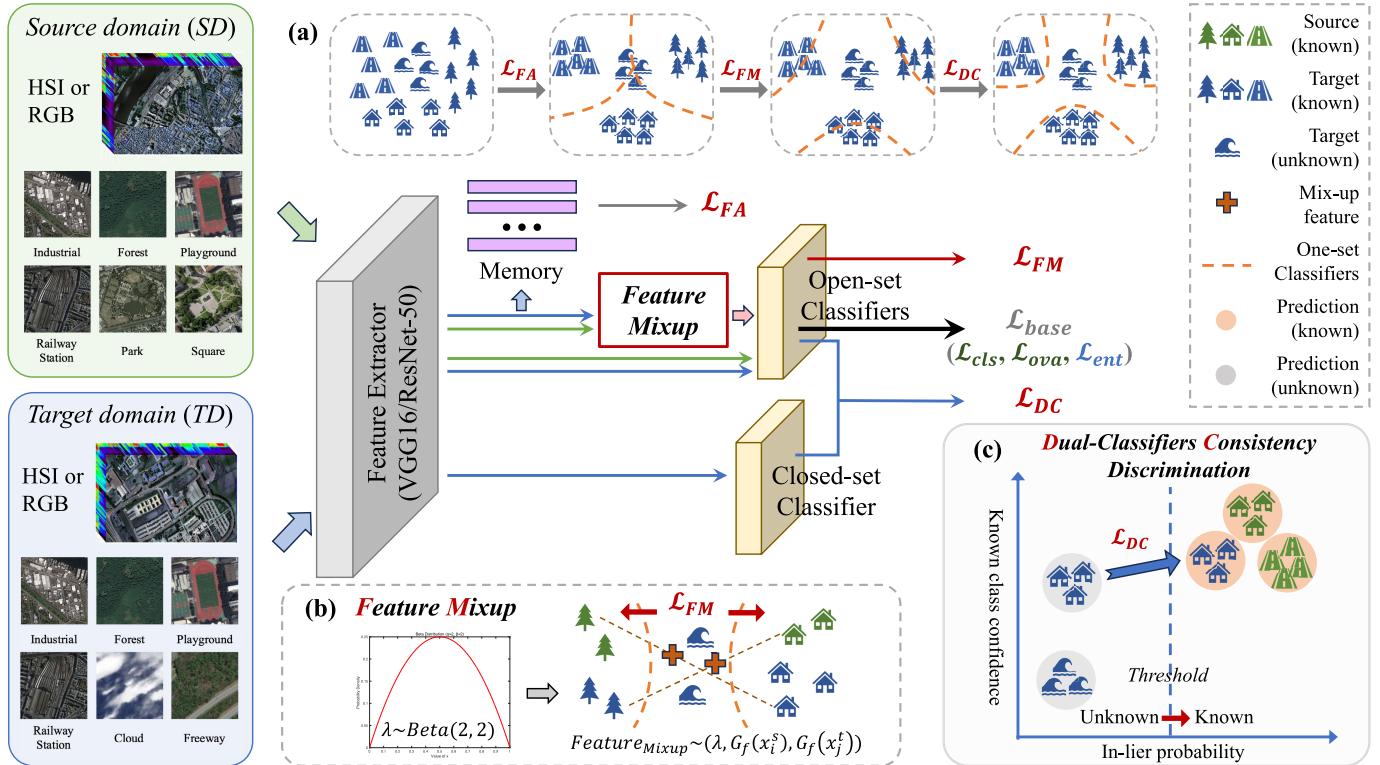


Fig. 2. Structure of our proposed DCmix. (a) Role of each module loss. (b) FM. The mixup features exhibit a low probability of belonging to a known class, thereby enhancing the identification of unknown classes. (c) Consistency discrimination. The minimization of \mathcal{L}_{dc} rectifies erroneous predictions of known-class samples as a consequence of high confidence.

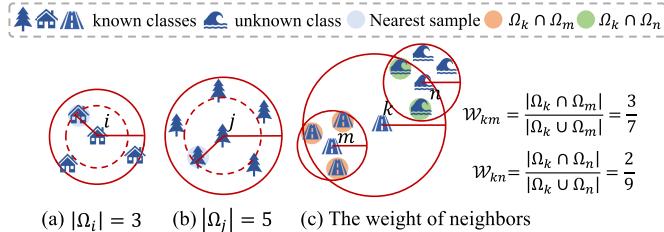


Fig. 3. Neighbor search and its confidence. (a) and (b) Adaptive neighborhood. (c) Weight of the sample neighbors, the neighborhood assigns increased confidence to neighbors that share more similarities.

our network is able to utilize various intermediate conditions generated by domain mixup, smoothing the classification of unseen data categories by open-set classifiers. The $\text{Mix}(x_i^s, x_j^t)$ is assigned to a known category if and only if x_i^s and x_j^t belong to the same known class, with the probability stated as

$$\begin{aligned} p\{(y_i^s = y_j^t) \in \mathcal{C}_s \cap \mathcal{C}_t\} &= \frac{\mathcal{S}\{(y_i^s = y_j^t) \in \mathcal{C}_s \cap \mathcal{C}_t\}}{\mathcal{S}\{y_i^s \in \mathcal{C}_s, y_j^t \in \mathcal{C}_t\}} \\ &= \frac{C_{|\mathcal{C}_s \cap \mathcal{C}_t|}^1}{C_{|\mathcal{C}_s|}^1 \cdot C_{|\mathcal{C}_t|}^1} = \frac{|\mathcal{C}_s \cap \mathcal{C}_t|}{|\mathcal{C}_s| \cdot |\mathcal{C}_t|} \end{aligned} \quad (7)$$

where \mathcal{S} is the number of atomic events. Due to $|\mathcal{C}_s \cap \mathcal{C}_t|$ is constrained by the minimum of $|\mathcal{C}_s|$ and $|\mathcal{C}_t|$, the value of $p\{(y_i^s = y_j^t) \in \mathcal{C}_s \cap \mathcal{C}_t\}$ represents a relatively small value. For instance, if $|\mathcal{C}_s \cap \mathcal{C}_t| = 13$, $|\mathcal{C}_s| = 30$, $|\mathcal{C}_t| = 21$, the $p\{(y_i^s = y_j^t) \in \mathcal{C}_s \cap \mathcal{C}_t\} \approx 0.021$. Therefore, $\text{Mix}(x_i^s, x_j^t)$ exhibits a low probability of belonging to a defined known category; the FM

is adept at simulating potential samples from unknown classes for the open-set classifier.

The loss based on the FM minimizes

$$\begin{aligned} \mathcal{L}_{FM}(x_i^s, y_i^s, x_j^t) &= L(\text{Mix}(x_i^s, x_j^t), \text{Mix}(y_i^s, \hat{y}_j)) \\ &= -\log(1 - P_o(y_i^s | \text{Mix}(x_i^s, x_j^t))) \end{aligned} \quad (8)$$

where $P_o(g|z) = \pi(z)_o$ represents the probability that the given feature z belongs to the class g as an in-lie, π is the softmax activation function. Through the minimization of \mathcal{L}_{FM} , the in-lie probability associated with simulated mixup instances are reduced, improving the classification performance for unknown-class samples.

D. Dual-Classifiers Consistency Discrimination

Although the FM can better identify unknown category samples, the incorporation of unknown samples in mixup procedures results in overcrowding of the decision boundaries for the open-set classifier. As the training progresses, the decision boundary for known classes progressively narrows, ultimately culminating in a degenerate state where all samples are erroneously labeled as unknown. In order to slow down the performance degradation of known-class recognition, we introduce the consistency constraint of the probability prediction distribution between closed-set and open-set classifiers.

Suppose that a known category sample x_i^t assigned the ground-truth label g is incorrectly identified as unknown. The closed-set classifier typically demonstrates high confidence in the maximum-probability known class g , yet its open-set classifier assigns it a low in-lie probability. In contrast,

Algorithm 1 Framework of DCmix for RS Images

Input: Source domain $\mathcal{D}^s = \{(\mathbf{x}_i^s, \mathbf{y}_i^s)\}_{i=1}^{n_s}$, target domain $\mathcal{D}^t = \{(\mathbf{x}_i^t)\}_{i=1}^{n_t}$, and trade-off parameter β, γ, φ .
Output: Well-trained feature extractor G_f and classifier G_c , predicted label $\hat{\mathbf{y}}^t$ for target domain.

- 1: Extract source and target samples and initialize the model.
- 2: Register target memory bank and neighbor matrix.
- 3: **while** not converged **do**
- 4: Calculate features, logits and probabilities for samples;
- 5: Update target feature bank;
- 6: Feature alignment, and calculate the \mathcal{W}_{ij} according to Equation (4);
- 7: **if** Corresponding loss weights exist **then**
- 8: Calculate source and target losses \mathcal{L}_{OVA} , neighbor loss \mathcal{L}_{FA} , FM loss \mathcal{L}_{FM} , and consistency loss \mathcal{L}_{dc} ;
- 9: Calculate total loss according to Equation (10), update metrics.
- 10: **end if**
- 11: Update the feature extractor G_f and classifier G_c .
- 12: **end while**
- 13: Obtain the predicted label $\hat{\mathbf{y}}^t$ for target domain.

unknown-class samples tend to exhibit lower known-class confidence. Consequently, we surmise that instances exhibiting discordant known class confidence and open-set in-lier probability are likely mislabeled known categories. This alignment of reduced known class confidence with a low open-set in-lier probability for unknown-class samples is visualized in Fig. 2(c). We compare the open-set in-lier probability from the corresponding open-set sub-classifier with a threshold of 0.5 to determine whether a sample belongs to a known or unknown class.

The dual-classifiers consistency discrimination based on known class confidence $P_c(g|x_i^t)$ and open-set in-lier probability $P_o(g|x_i^t)$ can be written as

$$\mathcal{L}_{dc}(x_i^t) = -\frac{1}{K_s} \sum_{g=1}^{K_s} \omega(x_i^t) P_c(g|x_i^t) \cdot P_o(g|x_i^t) \quad (9)$$

where K_s is the number of known classes, $\omega(x_i^t) = 1 - \max(P_c(g|x_i^t), P_o(g|x_i^t))$ is the weight of the sample x_i^t , and $P_c(g|x_i^t) = \pi(\mathcal{F}_\theta^t(x_i^t))_c$ represents the confidence that the given instance x_i^t belongs to the known class g . The partial derivative with regard to the open-set in-lier probability $P_o(g|x_i^t)$ assumes a negative value, adhering to the inequality. The $P_o(g|x_i^t)$ undergoes an increase that is directly proportional to the known class confidence $P_c(g|x_i^t)$ through the application of gradient descent. The minimization of \mathcal{L}_{dc} rectifies erroneous predictions of known-class samples as a consequence of high confidence. Conversely, for unknown samples, both $P_c(g|x_i^t)$ and $P_o(g|x_i^t)$ remain consistently low, minimizing the risk of misidentification as a known class.

Ultimately, we refine the open-set classifiers and their decision boundaries to effectively distinguish the known and unknown samples.

E. Overall Objective

We endeavor to optimize a suite of loss functions, specifically designed to address the complexities of RS images and the challenges associated with diverse source and target domains. The \mathcal{L}_{FA} enforces a similarity constraint on feature representations of reliable neighbors with high confidence, thereby minimizing intra-domain variations. By minimizing \mathcal{L}_{FM} , we reduce the in-lier probability of simulated mixup instances, enhancing classification performance for unknown-class samples. Furthermore, leveraging high closed-set confidence, \mathcal{L}_{dc} corrects the wrong classifications of known-class samples and optimizes the open-set classifier for distinguishing unknown classes.

Considering the aforementioned components, the overall objective can be expressed as

$$\mathcal{L}_{\text{overall}} = \mathcal{L}_{OVA} + \beta \mathcal{L}_{FA} + \gamma \mathcal{L}_{FM} + \varphi \mathcal{L}_{dc} \quad (10)$$

$$\mathcal{L}_{OVA} = \mathcal{L}_{ova} + \mathcal{L}_{cls} + 0.1 \mathcal{L}_{ent} \quad (11)$$

where β, γ, φ are hyperparameters that control the weights of different modules. The \mathcal{L}_{OVA} shares the identical loss function with the OVANet [25], specifically, \mathcal{L}_{ova} is the open-set classification loss, \mathcal{L}_{cls} refers to the cross-entropy loss for the closed-set classifier, and \mathcal{L}_{ent} signifies entropy of all the classifiers.

IV. DATASETS

A. HSI Datasets

Experiments use four HSI datasets, namely Houston2013, Houston2018, Pavia Center, and Pavia University, covering different geographical locations and spectral characteristics. Based on these datasets, we construct four HSI transfer tasks to evaluate our proposed DCmix method.

- 1) *Houston*: The Houston dataset is acquired over the University of Houston campus and the neighboring urban area. Due to differences in sensors and acquisition times, it is divided into **Houston2013** [50] and **Houston2018** [51]. We focus on the region (209 × 955 pixels) of these two scenes and investigate the commonalities in 48 spectral bands. As shown in Table I, these two scenes contain a total of seven categories. In our experiments, Non-residential buildings is selected as the source private class, and Road is selected as the target private class, with the remaining five categories serving as common classes. Fig. 4 provides a visualization of this dataset, presenting pseudo-color images alongside ground-truth maps.
- 2) *Pavia*: The Pavia dataset includes two subsets, Pavia Center (1096 × 715 pixels) and Pavia University (610 × 340 pixels). After discarding the last spectral band from Pavia University, both subsets use 102 spectral bands. As shown in Table II, both scenes share the same seven categories. In our experiments, Meadow is chosen as the source private class, and Bare soil is chosen as the target private class, with the remaining five categories serving as common classes. Fig. 5 provides a visualization of this dataset, presenting pseudo-color images alongside ground-truth maps.

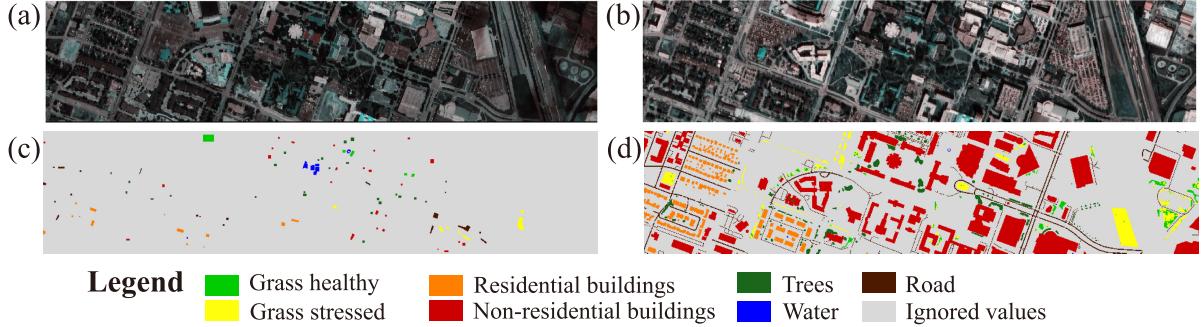


Fig. 4. Visualization for Houston task. (a) Houston2013 in pseudo-color. (b) Houston2018 in pseudo-color. (c) Houston2013 ground-truth map. (d) Houston2018 ground-truth map.

TABLE I
NUMBER OF SAMPLES FOR HOUSTON DATASETS

No.	Class	Number of Samples	
		Houston2013	Houston2018
1	Grass healthy	345	1353
2	Grass stressed	365	4888
3	Trees	365	2766
4	Water	285	22
5	Residential buildings	319	5347
6	Non-Residential buildings	408	32459
7	Road	443	6365
Total		2530	53200

TABLE II
NUMBER OF SAMPLES FOR PAVIA DATASETS

No.	Class	Number of Samples	
		Pavia Center	Pavia University
1	Tree	7598	3064
2	Asphalt	9248	6631
3	Brick	2685	3682
4	Bitumen	7287	1330
5	Shadow	2863	947
6	Meadow	3090	18649
7	Bare soil	6584	5029
Total		39355	39332

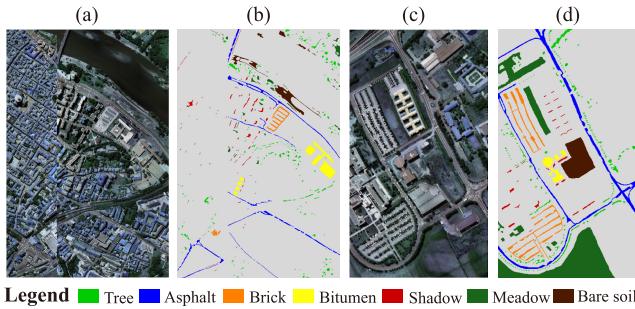


Fig. 5. Visualization for Pavia task. (a) Pavia University in pseudo-color. (b) Pavia Center in pseudo-color. (c) Pavia University ground-truth map. (d) Pavia Center ground-truth map.

B. RGB Datasets

Experiments involve constructing transfer learning tasks using three RS image datasets: AID (Aerial Image Dataset, A) [57], NWPU-RESISC45 (N) [58], and UC Merced (U) [59].

From these datasets, six transfer tasks can be established, with the specific domain categories for each task shown in Fig. 6. It is noteworthy that the labels in the two domains may not be identical. For example, A contains only the “Airport” category, whereas N distinguishes between “Airplane” and “Airport.” Additionally, since all categories in the U dataset are included in the N dataset, the $U \rightarrow N$ task is essentially an open-set DA task. Similarly, the $N \rightarrow U$ task is a partial DA task.

- 1) AID (Aerial Image Dataset, A) is an RGB dataset designed for RS image classification, including 30 scene categories (220~420 images each), totaling around 10 000 high-resolution (600×600 pixels) images. It covers various geographic and environmental conditions.
- 2) NWPU-RESISC45 (N) is a large-scale RGB dataset with 45 scene categories and approximately 700 images per category, totaling 31 500 images (256×256 pixels). It spans diverse scenes worldwide, including urban, rural, forest, and ocean environments, making it a valuable resource in RS image classification research.
- 3) UC Merced (U) is a smaller dataset with 21 scene categories, each with 100 images, totaling 2100 images (256×256 pixels). It covers various geographic environments in the United States and is often used for benchmark testing due to its high quality and well-defined categories.

V. EXPERIMENTAL RESULTS

A. Setup

To assess the performance of the introduced DCmix, we examine four HSI datasets and three RGB datasets. We utilize the PyTorch framework to instantiate and execute DCmix, leveraging the NVIDIA GeForce RTX 3090 for computations. Drawing inspiration from prior studies [23], [25], [54], [60], we adopt ResNet50, pre-trained on ImageNet [61], as the backbone network for RGB datasets. For HSI datasets, we adhere to the VGG16 architecture, similar to [24] and [62]. Patches in HSI datasets are standardized to a dimension of 12×12 , with labeling based on the central pixel’s category. The batch sizes for HSIs and RGB images are 64 and 36, respectively. The SGD optimizer with Nesterov momentum and an inverse scheduler is consistent with the setting detailed in [24]. For each hyperparameter, we set $\beta = 0.50$, $\gamma = 0.10$,

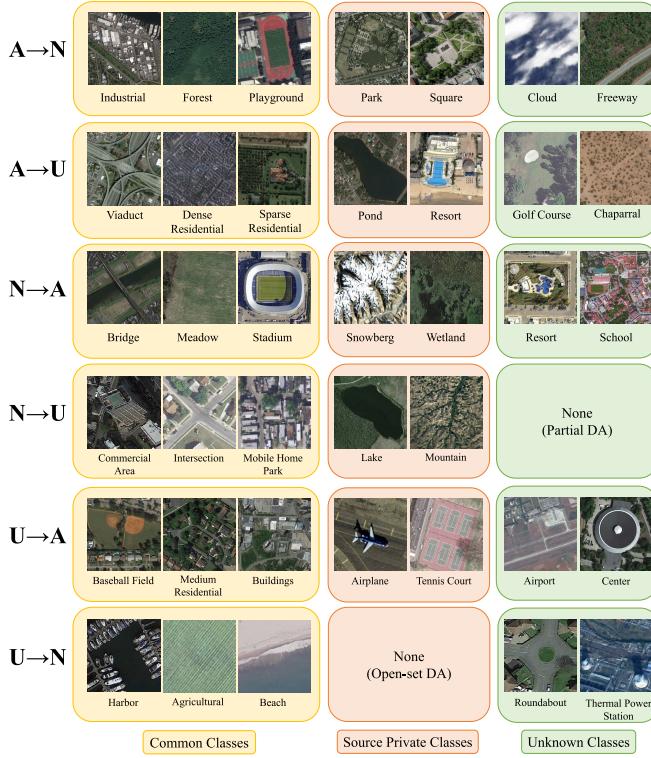


Fig. 6. Classes of image labels involved in each task and their examples. Since all the classes in \mathbf{U} are included in \mathbf{N} , the task of $\mathbf{U} \rightarrow \mathbf{N}$ is actually an open-set DA task. In the same way, the task of $\mathbf{N} \rightarrow \mathbf{U}$ is actually a partial DA task.

$\varphi = 0.16$, and the neighbor similarity ratio $\psi = 0.875$. In this study, we use the term overall accuracy (OA) to refer to the proportion of correctly classified samples across all categories. To more precisely evaluate the model's classification performance, we adhere to the evaluation metric H-score by Fu et al. [28], considering both common class (a_C) and target private class ($a_{\bar{C}_t}$), which is defined as

$$\text{H-score} = 2 \times \frac{a_C \times a_{\bar{C}_t}}{a_C + a_{\bar{C}_t}}. \quad (12)$$

B. Comparison Methods

To establish a comprehensive performance baseline, we incorporate a diverse set of SOTA DA algorithms for comparison, including domain-adversarial neural network (DANN) [52], joint adaptation networks (JAN) [63], conditional domain adversarial network (CDAN) [53], universal adaptation network (UAN) [21], domain adaptive neighborhood clustering via entropy optimization (DANCE) [54], calibrated multiple uncertainties (CMU) [28], OVA network (OVANet) [25], unified optimal transport framework for universal domain adaptation (UniOT) [29], universal domain adaptation for RS image scene classification (MA) [22], transferable vision transformer (TVT) [55], spectral alignment (SPA) [56], UniDA based on certainty, confidence, and consistency (C³DA) [23], and UniDA for cross-scene HSI classification (HyUniDA) [24]. Closed-set DA is difficult to accurately classify target private classes, according to its setting. Softmax regression is used to allocate diverse labels to the target domain, and

the label with the greatest likelihood is assigned to the target sample. When the probability of all known labels is below the threshold (0.9), we set it to unknown, i.e., target private. Additionally, we compare the baseline model (VGG16 for HSI and ResNet-50 for RGB), which only employs classification loss without DA techniques. Note that DCmix is a general framework independent of backbone architectures. The adoption of VGG16 and ResNet-50 in our experiments is solely for fair comparison with prior art. The performance improvements stem from our proposed consistency constraints and DA mechanisms, rather than the choice of backbones.

Tables III–VII provide the common classes accuracy, target private accuracy, OA, and H-score in the aforementioned methods for HSI and RGB target scenes, along with the classification accuracy for individual classes.

For HSI datasets, the proposed DCmix attains the highest H-score across all four scenarios, with the Houston2013→Houston2018 and Pavia Center→University scenarios also yielding the highest OA. Both DANN and CDAN, as standard DA methods, demonstrate higher OA than baseline, with the majority of H-scores also surpassing the baseline, except for CDAN on the Houston dataset, where its H-score is 0.1% lower than baseline. However, for target private class identification, the accuracy of these two methods is significantly lower than the other UniDA methods. DCmix enhances performance by minimizing \mathcal{L}_{FM} , which reduces the inlier probability of simulated mixup instances. SPAs perform well in the Houston2013→Houston2018 task but struggle in the reverse transfer task Houston2018→Houston2013, possibly due to a lack of generalization. Some methods exhibit overfitting, such as DANCE, UniOT, and MA, achieving an accuracy rate of 100% for the “water” category in the Houston scenario, which is significantly higher than the other categories. In the Houston task, DANCE reaches an accuracy of 94.13% for unknown categories but only 18.52% for known categories, possibly due to misclassifying known categories as unknown. By leveraging \mathcal{L}_{dc} , DCmix effectively corrects misclassifications for known class samples and optimizes the open-set classifier to distinguish unknown categories more accurately. Notably, for Pavia Center→University, our method achieves the highest accuracy of 64.69% and H-score of 68.91%, which are at least 2.45% and 8.14% higher than the other UniDA methods, respectively.

For RGB datasets, our method achieves the highest H-score except for U→N, where it is lower than C³DA. We use \mathcal{L}_{FM} to enforce similarity constraints on feature representations of reliable neighbors with high confidence, thus minimizing intra-domain variation. DCmix stands out in the evaluation of average OA and H-score, achieving the highest values of 63.09% and 62.09%, respectively. Among all comparison methods, the TVT achieved the highest OA of 61.07%, while the C³DA method, specifically designed for RGB images of RS, obtained the highest H-score of 58.44%. Compared with HSIs, RGB images have less band information. The gain of sample enhancement and mixup based on neighborhood relationship in DCmix is limited in a low-dimensional feature space, which may introduce additional noise.

Although the performance of DCmix may be similar to or slightly lower than that of SOTA methods in some categories, our method performs well in terms of stability and accuracy in

TABLE III
CLASSIFICATION ACCURACY METRICS, INCLUDING CLASS-SPECIFIC, COMMON, TARGET PRIVATE, OA, AND H-SCORE (%), FOR VARIOUS SOTA DA METHODS EVALUATED FROM **HOUSTON2013** TO THE TARGET SCENARIO **HOUSTON2018** (VGG16)

Class	Methods										
	VGG16	DANN [52]	CDAN [53]	UAN [21]	DANCE [54]	UniOT [29]	MA [22]	TVT [55]	SPA [56]	HyUniDA [24]	DCmix(Ours)
Grass healthy	2.07	10.50	9.31	29.34	54.84	83.74	82.93	49.22	30.45	72.26	51.52
Grass stressed	43.56	52.82	47.71	62.97	26.55	52.99	41.35	65.06	63.11	50.12	69.66
Trees	18.55	22.70	22.89	16.70	43.09	23.97	51.70	57.70	55.46	73.29	42.62
Water	90.91	90.91	90.91	77.27	100.00	95.45	100.00	86.36	95.45	63.55	90.91
Residential buildings	74.23	76.66	75.07	50.46	47.47	63.12	59.73	70.41	69.96	34.76	70.92
Common	46.32	51.97	49.56	46.27	40.29	54.13	54.18	64.18	61.16	50.89	63.25
Target private (Road)	55.18	52.98	50.98	71.39	74.27	58.65	57.38	37.64	65.15	81.67	68.77
Overall accuracy	49.04	52.28	50.00	53.98	50.72	55.52	55.16	56.03	62.39	60.33	64.94
H-score	50.36	52.47	50.26	56.15	52.24	56.30	55.73	47.45	63.10	62.71	65.89

TABLE IV
CLASSIFICATION ACCURACY METRICS, INCLUDING CLASS-SPECIFIC, COMMON, TARGET PRIVATE, OA, AND H-SCORE (%), FOR VARIOUS SOTA DA METHODS EVALUATED FROM PAVIA CENTER TO THE TARGET SCENARIO UNIVERSITY (VGG16)

Class	Methods										
	VGG16	DANN [52]	CDAN [53]	UAN [21]	DANCE [54]	UniOT [29]	MA [22]	TVT [55]	SPA [56]	HyUniDA [24]	DCmix(Ours)
Tree	41.16	68.18	60.38	54.86	70.40	69.61	62.37	30.58	74.41	77.74	70.56
Asphalt	51.30	59.46	57.83	43.76	79.73	20.99	45.42	69.55	61.53	95.48	61.08
Brick	11.62	7.03	5.13	25.94	29.44	2.04	39.73	0.00	5.68	0.22	51.33
Bitumen	0.00	0.15	0.08	0.23	6.77	4.51	0.08	0.00	0.00	20.83	11.43
Shadow	96.73	98.20	96.52	67.16	96.52	94.93	90.81	80.25	96.94	100.00	97.15
Common	38.37	46.14	43.37	39.46	60.89	29.12	46.29	40.30	47.83	63.53	58.60
Target private (Bare soil)	16.62	45.16	26.39	59.14	53.47	58.12	56.25	31.90	76.75	58.24	83.62
Overall accuracy	33.09	45.90	39.24	44.24	59.09	36.17	48.72	38.26	54.86	62.24	64.69
H-score	23.20	45.64	32.81	47.33	56.94	38.80	50.79	35.61	58.93	60.77	68.91

TABLE V
CLASSIFICATION ACCURACY METRICS, INCLUDING CLASS-SPECIFIC, COMMON, TARGET PRIVATE, OA, AND H-SCORE (%), FOR VARIOUS SOTA DA METHODS EVALUATED FROM **HOUSTON2018** TO THE TARGET SCENARIO **HOUSTON2013** (VGG16)

Class	Methods										
	VGG16	DANN [52]	CDAN [53]	UAN [21]	DANCE [54]	UniOT [29]	MA [22]	TVT [55]	SPA [56]	HyUniDA [24]	DCmix(Ours)
Grass healthy	99.42	99.42	99.42	6.37	0.00	44.05	5.22	4.93	47.25	88.99	89.56
Grass stressed	17.53	15.62	1.64	49.86	0.00	27.80	29.59	27.95	61.37	40.00	38.63
Trees	14.52	7.40	6.30	6.03	0.27	36.84	32.05	97.53	46.58	10.68	10.14
Water	1.75	1.75	1.75	68.42	31.93	100.00	38.60	9.82	0.00	7.37	5.26
Residential buildings	45.45	47.02	47.33	13.79	68.65	38.69	52.04	62.70	42.95	48.28	67.40
Common	36.27	34.66	31.45	27.70	18.52	35.23	30.91	42.35	41.33	39.73	42.70
Target private (Road)	31.83	43.34	32.05	88.04	94.13	80.63	43.87	35.78	33.09	62.08	67.27
Overall accuracy	35.34	36.48	31.57	40.29	34.31	49.16	33.45	41.06	39.72	44.39	47.83
H-score	33.91	38.52	31.74	42.13	30.95	49.04	36.26	38.79	36.75	48.45	52.24

TABLE VI
CLASSIFICATION ACCURACY METRICS, INCLUDING CLASS-SPECIFIC, COMMON, TARGET PRIVATE, OA, AND H-SCORE (%), FOR VARIOUS SOTA DA METHODS EVALUATED FROM PAVIA UNIVERSITY TO THE TARGET SCENARIO CENTER(VGG16)

Class	Methods										
	VGG16	DANN [52]	CDAN [53]	UAN [21]	DANCE [54]	UniOT [29]	MA [22]	TVT [55]	SPA [56]	HyUniDA [24]	DCmix(Ours)
Tree	78.43	64.53	75.45	42.41	93.71	60.28	80.35	99.08	82.59	97.25	81.61
Asphalt	31.15	12.55	34.70	31.54	12.74	22.42	35.15	72.89	72.23	6.70	49.31
Brick	79.55	47.45	48.79	54.41	52.96	16.76	44.58	5.59	19.44	3.02	59.96
Bitumen	82.85	83.64	77.04	0.75	83.85	17.83	13.71	59.28	6.42	98.52	12.08
Shadow	81.52	78.87	80.13	24.55	90.36	72.13	82.36	93.61	65.63	95.91	95.98
Common	65.18	52.86	61.18	28.16	62.05	35.26	46.86	72.16	54.00	60.69	53.90
Target private (Bare soil)	40.35	52.44	50.88	81.18	41.08	57.78	50.13	43.07	51.10	48.91	71.39
Overall accuracy	60.67	52.79	59.31	37.79	58.24	39.35	47.17	69.42	53.73	58.56	57.08
H-score	49.85	52.65	55.56	41.81	49.44	43.80	48.44	53.95	52.51	54.17	61.42

the overall task. Especially in different cross-domain scenarios, DCmix does not overfit and maintains strong generalization ability through effective module synergy.

Figs. 7 and 8 show the classification maps of each comparison algorithm for target scenarios Houston2018 and Pavia

University, respectively. In the depicted maps, the marked pixels represent the corresponding model-predicted categories, while the unmarked pixels indicate background or ignored points. From these maps, it is observable that our proposed DCmix method demonstrates lower noise and improved

TABLE VII
OA AND H-SCORE (%) ON OUR COLLECTED DATASET FOR UNIDA SCENARIOS (RESNET-50)

Method	A → N		A → U		N → A		N → U		U → A		U → N		Avg	
	OA	H-score	OA	H-score	OA	H-score	OA	OA	H-score	OA	H-score	OA	OA	H-score
ResNet-50 [61]	67.66	35.23	49.46	44.74	73.09	45.64	41.14	52.75	23.00	40.33	43.98	54.07	38.52	
DANN [52]	47.20	48.87	43.61	30.05	66.29	44.90	37.71	45.57	41.14	67.62	52.19	51.33	43.43	
JAN [63]	55.82	30.90	48.34	36.47	65.87	45.53	43.10	51.88	38.34	55.52	60.38	53.42	42.32	
CDAN [53]	48.40	50.59	30.29	35.38	61.72	50.14	39.62	50.15	23.49	49.67	61.28	46.64	44.18	
UAN [21]	60.83	20.96	37.42	27.93	31.93	6.90	20.52	49.89	17.97	59.48	71.29	43.35	29.01	
CMU [28]	56.49	40.41	42.29	36.45	69.76	53.14	54.86	53.73	35.62	64.62	71.38	56.96	47.40	
DANCE [54]	57.84	17.51	39.14	46.61	64.37	22.58	54.52	54.29	52.90	63.14	71.08	55.55	42.13	
OVANet [25]	43.73	55.12	36.77	45.11	47.65	59.17	37.57	42.64	50.44	54.52	58.61	43.81	53.32	
UniOT [29]	57.67	63.22	33.31	42.05	73.46	61.36	52.14	54.55	55.31	58.08	58.65	54.87	56.12	
MA [22]	64.97	44.51	45.31	46.07	64.92	52.34	52.72	54.64	57.76	56.03	70.45	56.88	54.23	
TVT [55]	61.61	60.63	54.43	52.32	68.92	46.67	72.29	55.64	52.21	53.51	52.61	61.07	52.89	
SPA [56]	60.63	60.00	52.48	50.90	69.86	63.70	72.33	54.11	53.85	56.72	53.70	61.02	56.43	
C ³ DA [23]	60.13	59.68	51.57	45.70	65.74	64.07	55.33	55.97	50.91	70.71	73.16	58.81	58.44	
DCmix(Ours)	66.29	63.59	49.45	56.07	65.96	68.29	70.33	55.41	60.95	71.07	64.54	63.09	62.69	

accuracy in identifying certain categories. As shown in Fig. 7, our method more accurately identifies trees and differentiates between healthy grass and stressed grass. Similarly, as depicted in Fig. 8, our method excels in identifying bare soil, which is a target private class. DCmix demonstrates exceptional noise reduction and accuracy improvement, surpassing other comparative DA methods. The proposed method effectively captures categories that appear only in the target domain, showcasing its robust classification capabilities. This ensures alignment between the maps produced from predictions and the actual ground truth maps. Overall, DCmix proves to be a robust and efficient method for HSI and RGB image classification tasks.

We performed a detailed analysis of the computational complexity of our method compared to several SOTA approaches, as summarized in Table VIII. The table reports training speed in frames per second (fps) and the number of parameters (in millions, M). The fps metric provides a direct and interpretable measure of how many images a model can process per second during inference. Our method achieves a training speed of 27.78 fps, which is comparable to other competitive approaches such as OVANet, MA, and HyUniDA. While our method does not outpace the fastest method (C³DA, 29.11 fps), it maintains a competitive performance among existing approaches. Regarding parameters, our method requires 23.57 M, which is either on par with or lower than several other methods, underscoring its computational efficiency without introducing significant overhead.

C. Ablation Studies

For validation and assessment of the DCmix's efficacy, we conducted a series of ablation experiments aimed at gaining in-depth insights into the contributions of individual components to overall performance.

We conducted experiments using various combinations of loss modules. For HSI datasets, as shown in Table IX, when no loss modules are used, the baseline model achieves an accuracy of 54.26%~55.58% and an H-score of 28.13%~30.35%. Adding only \mathcal{L}_{FA} significantly increases the accuracy of known classes to 67.23% and 77.37% (OA gains of 1.48% and 3.63%), while H-score increases marginally to 30.81% and

TABLE VIII
COMPARISON OF COMPUTATIONAL COMPLEXITY BETWEEN THE METHODS

Method	Training Speed (fps)	Parameters (M)
ResNet-50 [61]	48.03	25.56
UAN [21]	24.87	28.71
CMU [28]	26.34	25.38
DANCE [54]	25.94	23.53
OVANet [25]	27.01	23.53
UniOT [29]	26.05	28.27
MA [22]	27.03	27.03
C ³ DA [23]	29.11	26.70
SPA [56]	24.75	24.04
HyUniDA [24]	27.15	24.04
Ours	27.78	23.57

29.46%, indicating neighborhood alignment primarily benefits known-class discrimination. Combining \mathcal{L}_{FM} with \mathcal{L}_{FA} significantly enhances the recognition of unknown classes, as evidenced by the H-score increasing from 30.35% to 61.73% on Pavia and from 28.13% to 52.81% on Houston. The introduction of the \mathcal{L}_{FM} reduces the inlier probability associated with simulated mixup instances, thereby improving the classification performance for unknown-class samples. Enabling only \mathcal{L}_{FM} dramatically boosts unknown class to 74.69% (Pavia) and 55.26% (Houston) but drops known class performance and OA, showing that \mathcal{L}_{FM} is prone to introduce noise and blur the decision boundary in the absence of neighborhood information. Building on this, adding the \mathcal{L}_{dc} further enhances DCmix's performance by rectifying overconfident predictions for known-class samples. OA and H-score increase by 6.90%~8.13% and 7.18%~13.08%, respectively. Using a combination of \mathcal{L}_{FA} and \mathcal{L}_{dc} or \mathcal{L}_{FM} and \mathcal{L}_{dc} results in limited performance improvements. The best performance for all tasks is achieved when all three loss terms are activated, indicating that they are complementary and beneficial for UniDA scenarios.

Table IX demonstrates that our proposed modules deliver substantial performance gains with minimal computational overhead. The baseline model requires 23.65 s/epoch and 2.37 GB of memory, whereas the full DCmix configuration

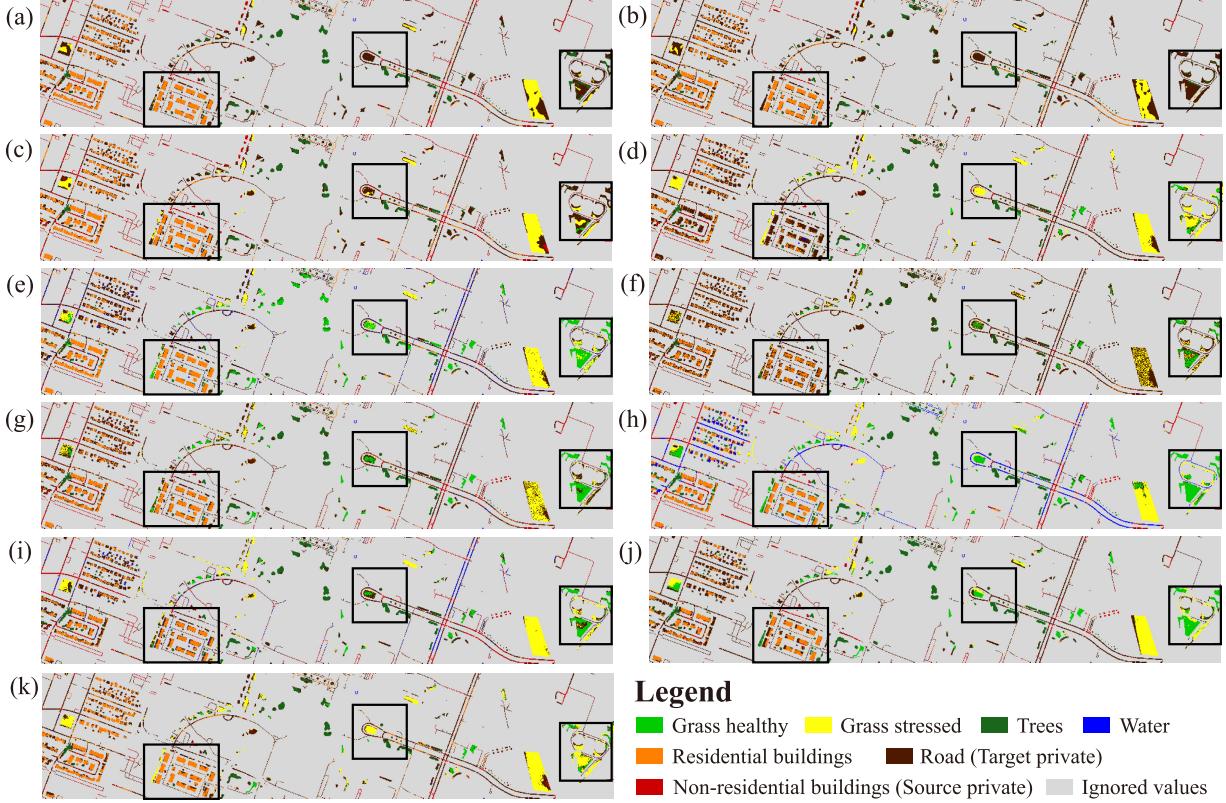


Fig. 7. Classification maps from source domain for the target scenario **Houston2018** produced by multiple methods, encompassing (a) VGG16, (b) DANN, (c) CDAN, (d) UAN, (e) DANCE, (f) UniOT, (g) MA, (h) TVT, (i) SPA, (j) HyUniDA, and (k) DCmix(Ours).

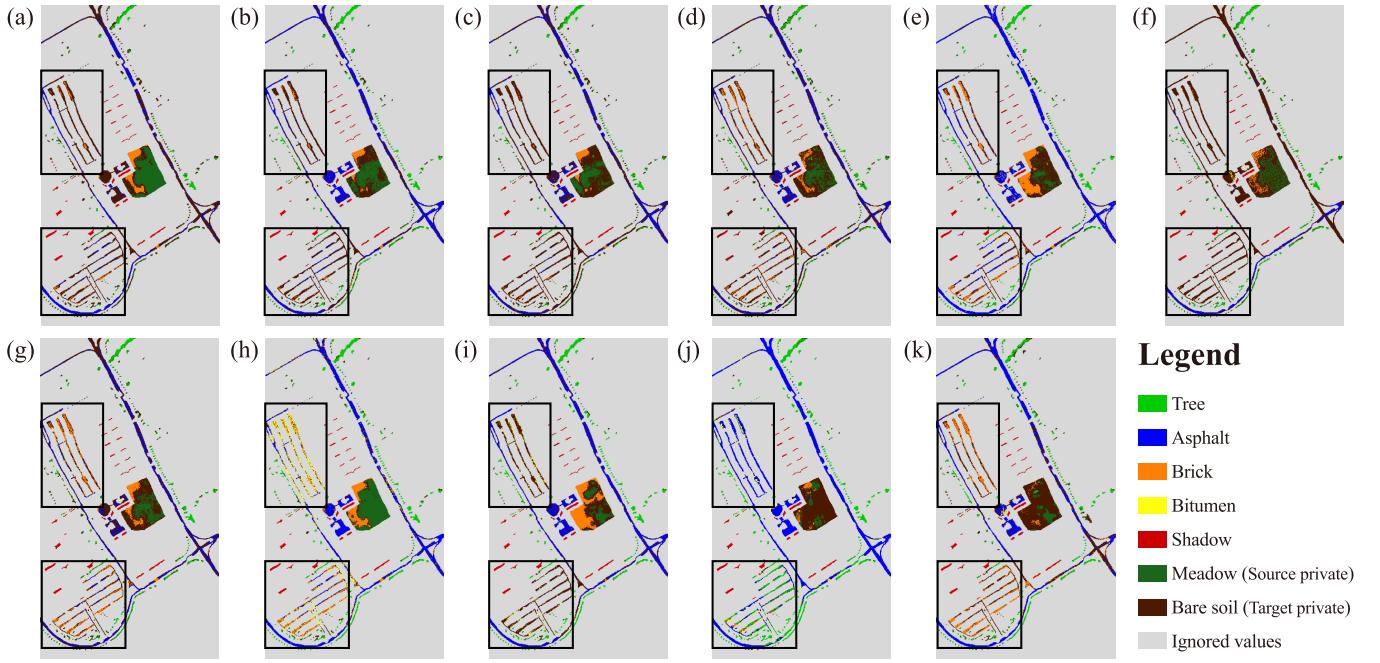


Fig. 8. Classification maps from source domain for the target scenario **Pavia University** produced by multiple methods, encompassing (a) VGG16, (b) DANN, (c) CDAN, (d) UAN, (e) DANCE, (f) UniOT, (g) MA, (h) TVT, (i) SPA, (j) HyUniDA, and (k) DCmix(Ours).

$(\mathcal{L}_{FA} + \mathcal{L}_{FM} + \mathcal{L}_{dc})$ takes 24.05 s and 2.51 GB. All module additions incur less than 1.7% extra training time and under

6% additional memory, while the H-score increases dramatically from 30.35% to 68.91%.

TABLE IX
OA, H-SCORE (%), TRAINING TIME OF ONE EPOCH (s) AND MEMORY CONSUMPTION (GB) FOR DIFFERENT LOSS COMPONENTS ON THE HSI DATASETS (VGG16)

\mathcal{L}_{FA}	\mathcal{L}_{FM}	\mathcal{L}_{DC}	Pavia Center → University				Houston2013 → Houston2018				Training time	Peak memory
			Known	Unknown	OA	H-score	Known	Unknown	OA	H-score		
✓	✓	✓	65.34	19.77	54.26	30.35	72.45	17.45	55.58	28.13	23.65±0.59	2.3738
			67.23	19.98	55.74	30.81	77.37	18.19	59.21	29.46	23.92±0.43	2.5123
	✓	✓	41.39	74.69	49.49	53.26	42.84	55.26	46.65	48.26	23.75±0.50	2.3741
			57.85	23.82	49.58	33.75	66.27	21.45	52.51	32.40	23.92±0.61	2.3737
	✓	✓	51.88	76.20	57.79	61.73	61.46	46.30	56.81	52.81	23.70±0.64	2.5126
✓	✓	✓	69.62	26.35	59.10	38.22	66.99	25.26	54.18	36.69	23.78±0.38	2.5127
✓	✓	✓	46.79	81.53	55.23	59.45	48.80	57.38	51.43	52.74	23.77±0.46	2.3767
✓	✓	✓	58.60	83.62	64.69	68.91	63.25	68.77	64.94	65.89	24.05±0.47	2.5128

This also applies to RGB datasets. As indicated in Table X, the model's overall performance decreases when any single module is removed. After removing the \mathcal{L}_{FA} , there is a significant drop in the OA, indicating that this module plays a key role in the recognition of known classes. Notably, when \mathcal{L}_{FM} is excluded in the A→U scenario, the H-score drops from 56.07% to 26.73%. The introduction of the \mathcal{L}_{FM} module is crucial for the recognition of unknown classes. This comprehensive analysis of these loss modules highlights their importance in enhancing model performance, showing that the combined use of the three loss modules is crucial for improving model performance.

D. Sensitivity to Hyperparameters

We carried out a sensitivity analysis on four crucial hyperparameters, i.e., β , γ , φ , and ψ . The impact of varying β and γ values on the OA and H-score metrics for the Pavia and Houston datasets is presented in Figs. 9 and 10 provides an analysis of how variations in the parameter φ affect the model's performance, specifically when evaluated on the RGB dataset. Fig. 11 illustrates the experiments on varying neighbor similarity ratio ψ . We evaluated the impact of different β values within the range of 0.4–0.6 on OA and H-score. As depicted in Fig. 9(a), when β is less than 0.5, both accuracy and H-score increase with the increase of β . However, when β exceeds 0.5, both metrics sharply decrease in the Pavia and Houston datasets. Similarly, in Fig. 9(b), a γ value of 0.1 stands out as a peak, significantly outperforming other settings, indicating that \mathcal{L}_{FM} functions best at this weight. For φ , we conducted a sensitivity analysis on six tasks within the RGB dataset. As illustrated in Fig. 10(a) and (b), we examined the changes in OA and H-score as φ varied from 0.12 to 0.20. For most tasks, φ equals 0.16 is the optimal value for both metrics. However, for certain tasks, such as the N→U scenario, both accuracy and H-score are lower at $\varphi = 0.16$ compared to other settings. Regarding the parameter ψ , we varied it from 0.800 to 0.990. When ψ equals 0.875, both the Pavia dataset and the A→U task achieve the highest H-score. For accuracy, the maximum value for the Pavia dataset also occurs at $\psi = 0.875$, but for the A→U task, the accuracy at this value is lower than most of the other settings. Based on our findings, we set these hyperparameters to the following optimal values: $\beta = 0.50$, $\gamma = 0.10$, $\varphi = 0.16$, and the neighbor similarity ratio $\psi = 0.875$.

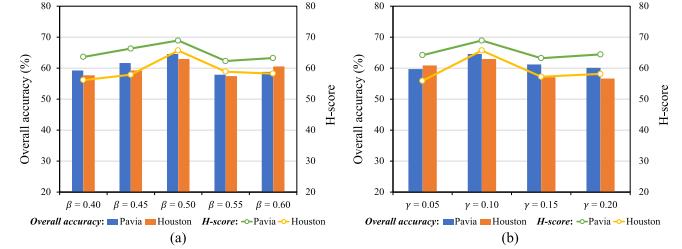


Fig. 9. Sensitivity of model performance to hyperparameters on the HSI datasets (VGG16). (a) Effect of β on OA and H-score. (b) Effect of γ on OA and H-score.

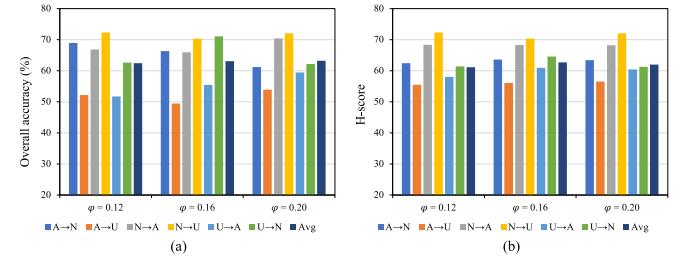


Fig. 10. Sensitivity of model performance to hyperparameter φ on RGB datasets (ResNet-50). (a) OA varies with different φ . (b) H-score varies with different φ .

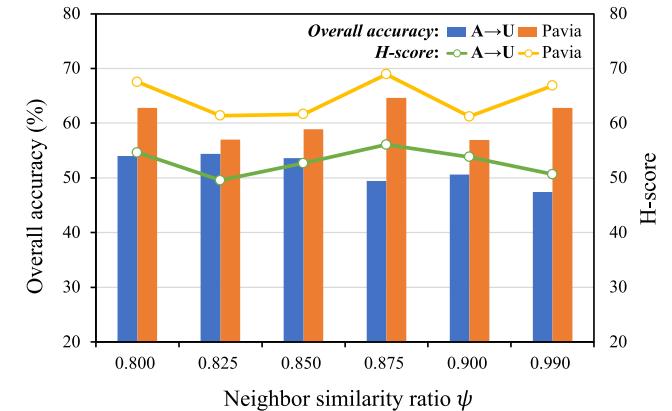


Fig. 11. Sensitivity to neighbor similarity ratio ψ on RGB datasets ($A \rightarrow U$) and HSI datasets (Pavia Center → University).

E. Robustness in Realistic UniDA

1) t-SNE Visualizations: The t-distributed stochastic neighbor embedding (t-SNE) [64] provides a 2-D representation of

TABLE X
OA AND H-SCORE (%) FOR DIFFERENT LOSS COMPONENTS ON THE RGB DATASETS (RESNET-50)

Method	A → N		A → U		N → A		N → U		U → A		U → N		Avg	
	OA	H-score	OA	H-score	OA	H-score	OA	OA	H-score	OA	H-score	OA	OA	H-score
w/o \mathcal{L}_{DC}	66.10	60.33	52.01	49.09	68.69	59.14	61.86	54.86	57.02	55.04	58.08	59.76	56.73	
w/o \mathcal{L}_{FM}	54.73	56.27	55.03	26.73	62.02	60.68	64.76	53.58	47.50	58.97	49.48	58.18	48.13	
w/o \mathcal{L}_{FA}	60.50	62.94	48.13	52.06	72.69	67.61	62.14	45.32	52.51	50.03	58.88	56.47	58.80	
Ours (full)	66.29	63.59	49.45	56.07	65.96	68.29	70.33	55.41	60.95	71.07	64.54	63.09	62.69	

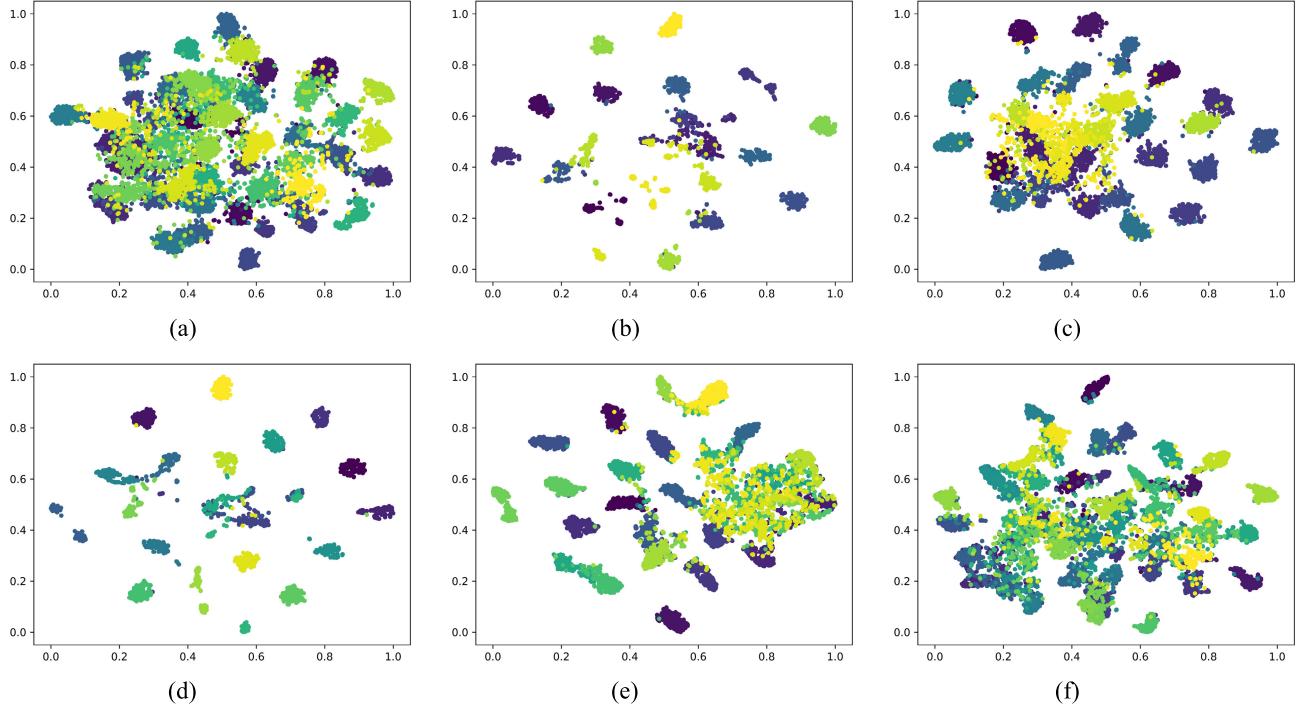


Fig. 12. t-SNE visualization of features for six transfer tasks (ResNet-50). Yellow plots are “unknown” samples, and others are “known” samples. (a) A→N. (b) A→U. (c) N→A. (d) N→U. (e) U→A. (f) U→N.

the high-dimensional feature space, showcasing the separation and clustering of known and unknown classes. To demonstrate the feature transferability of our DCmix, we illustrate the visual representations of network features obtained from the final convolutional layer for six diverse transfer learning tasks in Fig. 12. Moreover, Fig. 13 shows the t-SNE feature visualization of the extracted features for each incremental module in the A→U task. Notably, the clear separation between known and unknown samples demonstrates the network’s capacity to distinguish between these classes effectively in Fig. 12(b)–(d). The clustering pattern observed among known samples also indicates that our method has effectively minimized intra-domain variations and substantially enhanced the generalization ability of the features. Features remain well-organized and distinct even when transferring knowledge across significantly different domains. This validates our approach for smoother decision boundaries and more compact feature representations, contributing to improved classification performance in RS image analysis.

2) *Different Class Splits:* In a real cross-domain HSI classification scenario, the sample class of the target domain is entirely unknown, thus resulting in a multitude of class splits. To assess the robustness of our proposed DCmix in realistic UniDA, we carry out experiments with the task A→N,

TABLE XI
ROBUSTNESS IN REALISTIC UNIDA WITH DIVERSE RATIOS OF COMMON CLASSES FOR THE TASK **A → N**. DIFFERENT CLASS SPLITS $\mathcal{C}/\bar{\mathcal{C}}_s/\bar{\mathcal{C}}_t$ REVEAL DIVERSE RATIOS OF COMMON CLASSES

	Class split					
	9/7/7	10/6/7	11/6/6	12/6/5	13/5/5	14/5/4
Known	66.95	71.64	74.95	69.88	77.05	75.15
Unknown	70.47	60.57	61.07	66.00	62.97	66.39
Overall accuracy	67.30	70.64	73.80	69.58	76.05	74.56
H-score	68.67	65.64	67.30	67.89	69.30	70.50

considering diverse proportions of common classes, as shown in Table XI. It is observed that DCmix achieves high OA and H-score across different class splits, demonstrating robust generalization capabilities toward shared classes. The OA ranges from 67.30% to 76.05%, while the H-score varies from 65.64% to 70.50%. The consistent results indicate that DCmix effectively balances the performance between common and target private classes under varying proportions. Furthermore, the method shows effectiveness in feature learning for the target domain, enabling accurate distinction of target-specific samples even when they are not present in the source domain, with accuracies ranging from 60.57% to 70.47%. In summary, DCmix is not sensitive to the variation of class splits.

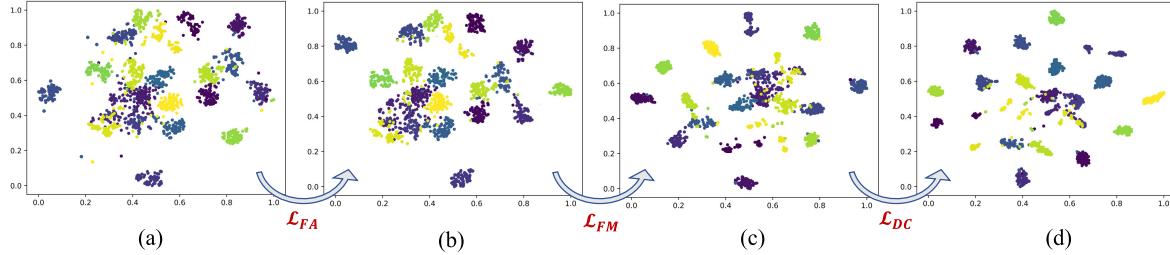


Fig. 13. t-SNE feature alignment for the A-U task under different setups. (a) Base setup, (b) with \mathcal{L}_{FA} , (c) with $\mathcal{L}_{FA} + \mathcal{L}_{FM}$, and (d) with $\mathcal{L}_{FA} + \mathcal{L}_{FM} + \mathcal{L}_{dc}$.

VI. CONCLUSION

In this study, we confront the significant issue of UniDA in RS image classification by introducing DCmix, a dual-classifier network. The architecture integrates both closed-set and open-set classifiers, marking the first application of such a dual-classifier approach to the UniDA task in the RS community. Our method improves the identification accuracy of unknown sample classes by leveraging the complementary strengths of these classifiers. We enhance the generalization capability of features in the target domain by using the neighborhood relations of samples. The feature alignment module includes reliable neighbor search, similarity maximization, and confidence measurement, thereby fostering more robust and adaptable feature representations. Additionally, we implement a cross-domain FM scheme guided by the consistency discrimination of the dual classifiers. The FM scheme not only produces smoother decision boundaries but also simplifies hidden layer representations by explicitly simulating intermediate states of unknown classes. Comprehensive experiments performed on multiple hyperspectral and RGB datasets have demonstrated that our approach attains SOTA performance in the UniDA scenario for RS image classification, validating its effectiveness and significant contributions to the field.

APPENDIX

PROOF OF MITIGATION FOR SAMPLE IMBALANCE

The k_i and n_i correspond to the neighbor count and total samples of the i th class, respectively, while k_j and n_j represent the same for the j th class. We assume a disparity in the data, where the i th class has a higher sample count ($n_i > n_j$) compared to the j th class. The relative neighbor ratio between the i th and the j th classes is denoted as

$$\mathcal{R}_{ij} = \frac{k_i}{n_i} \div \frac{k_j}{n_j} = \frac{k_i}{k_j} \cdot \frac{n_j}{n_i}. \quad (13)$$

If $\mathcal{R}_{ij} = 1$, it implies a balanced distribution of neighbors across the two classes. The k_i and k_j are both assigned the same value in K -nearest neighbors; the relative neighbor ratio can be condensed to $\mathcal{R}_{ij}^\dagger = n_j/n_i$. The K -nearest neighbors approach has a preference for sampling the j th class data more often.

During the neighbor search, the number of neighbor samples chosen for each category is proportional to its sample size, i.e., $k_i \propto n_i$, $k_j \propto n_j$. Owing to $n_i > n_j$, we may logically postulate that $k_i > k_j$, resulting in the inequality

$$\mathcal{R}_{ij} = \frac{k_i}{k_j} \cdot \frac{n_j}{n_i} > \frac{k}{k} \cdot \frac{n_j}{n_i} > 0. \quad (14)$$

The quantity of neighbors chosen for the i th class is scaled up by the factor $k_i/k_j > 1$. The \mathcal{R}_{ij} approaches 1 more closely than \mathcal{R}_{ij}^\dagger , leading to a more balanced neighbor search and subsequently reducing the data imbalance.

REFERENCES

- [1] Y. Wang, Y. Sun, X. Cao, Y. Wang, W. Zhang, and X. Cheng, "A review of regional and global scale land use/land cover (LULC) mapping products generated from satellite remote sensing," *ISPRS J. Photogramm. Remote Sens.*, vol. 206, pp. 311–334, Dec. 2023.
- [2] J. Zheng et al., "Surveying coconut trees using high-resolution satellite imagery in remote atolls of the Pacific ocean," *Remote Sens. Environ.*, vol. 287, Mar. 2023, Art. no. 113485.
- [3] J. Zheng, S. Yuan, W. Li, H. Fu, L. Yu, and J. Huang, "A review of individual tree crown detection and delineation from optical remote sensing images: Current progress and future," *IEEE Geosci. Remote Sens. Mag.*, vol. 13, no. 1, pp. 209–236, Mar. 2025.
- [4] J. Im, H. Park, and W. Takeuchi, "Advances in remote sensing-based disaster monitoring and assessment," *Remote Sens.*, vol. 11, no. 18, p. 2181, Sep. 2019.
- [5] Q. Zhu et al., "A spectral-spatial-dependent global learning framework for insufficient and imbalanced hyperspectral image classification," *IEEE Trans. Cybern.*, vol. 52, no. 11, pp. 11709–11723, Nov. 2022.
- [6] J. Zheng et al., "A two-stage adaptation network (TSAN) for remote sensing scene classification in single-source-mixed-multiple-target domain adaptation (S^2M^2T DA) scenarios," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2021, Art. no. 5609213.
- [7] Y. Chen, X. Zhao, and X. Jia, "Spectral-spatial classification of hyperspectral data based on deep belief network," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 6, pp. 2381–2392, Jun. 2015.
- [8] X. Wu, D. Hong, and J. Chanussot, "Convolutional neural networks for multimodal remote sensing data classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5517010.
- [9] H. M. Fathima, M. Lakshmi, P. Pooja, A. Anitha, and K. Kalaiselvi, "Convolutional neural network for remote sensing classification," in *Proc. AIP Conf.*, 2023, vol. 2831, no. 1, Art. no. 020002.
- [10] A. Ma, N. Yu, Z. Zheng, Y. Zhong, and L. Zhang, "A supervised progressive growing generative adversarial network for remote sensing image scene classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5618818.
- [11] D. Li and R. Zhang, "Ensemble stacked auto-encoder classification on LiDAR remote sensing images," *J. Indian Soc. Remote Sens.*, vol. 46, no. 4, pp. 597–604, Apr. 2018.
- [12] S. K. Roy, A. Deria, D. Hong, B. Rasti, A. Plaza, and J. Chanussot, "Multimodal fusion transformer for remote sensing image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 5515620.
- [13] J. Zheng et al., "Cross-regional oil palm tree counting and detection via a multi-level attention domain adaptation network," *ISPRS J. Photogramm. Remote Sens.*, vol. 167, pp. 154–177, Sep. 2020.
- [14] Y. Shi, L. Du, C. Li, Y. Guo, and Y. Du, "Unsupervised domain adaptation for SAR target classification based on domain-and class-level alignment: From simulated to real data," *ISPRS J. Photogramm. Remote Sens.*, vol. 207, pp. 1–13, Jan. 2024.
- [15] W. Chen, Y. Wen, J. Zheng, J. Huang, and H. Fu, "BAN: A universal paradigm for cross-scene classification under noisy annotations from RGB and hyperspectral remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 63, 2025, Art. no. 5610213.

- [16] J. Zheng, Y. Zhao, W. Wu, M. Chen, W. Li, and H. Fu, "Partial domain adaptation for scene classification from remote sensing imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 5601317.
- [17] L. Kong et al., "Partial identifiability for domain adaptation," 2023, *arXiv:2306.06510*.
- [18] Y. Ma, Z. Yang, Q. Huang, and Z. Zhang, "Improving the transferability of deep learning models for crop yield prediction: A partial domain adaptation approach," *Remote Sens.*, vol. 15, no. 18, p. 4562, Sep. 2023.
- [19] Y. Xu, L. Chen, L. Duan, I. W. Tsang, and J. Luo, "Open set domain adaptation with soft unknown-class rejection," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 3, pp. 1601–1612, Mar. 2023.
- [20] J. Zheng et al., "Open-set domain adaptation for scene classification using multi-adversarial learning," *ISPRS J. Photogramm. Remote Sens.*, vol. 208, pp. 245–260, Feb. 2024.
- [21] K. You, M. Long, Z. Cao, J. Wang, and M. I. Jordan, "Universal domain adaptation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2019, pp. 2720–2729.
- [22] Q. Xu, Y. Shi, X. Yuan, and X. X. Zhu, "Universal domain adaptation for remote sensing image scene classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 4700515.
- [23] J. Guo et al., "C³DA: A universal domain adaptation method for scene classification from remote sensing imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 21, 2024, Art. no. 6006705.
- [24] Q. Li, Y. Wen, J. Zheng, Y. Zhang, and H. Fu, "HyUniDA: Breaking label set constraints for universal domain adaptation in cross-scene hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 62, 2024, Art. no. 5518415.
- [25] K. Saito and K. Saenko, "OVANet: One-vs-all network for universal domain adaptation," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2021, pp. 9000–9009.
- [26] L. Chen, Y. Lou, J. He, T. Bai, and M. Deng, "Geometric anchor correspondence mining with uncertainty modeling for universal domain adaptation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2022, pp. 16134–16143.
- [27] Y. Lu, M. Shen, A. J. Ma, X. Xie, and J.-H. Lai, "MLNet: Mutual learning network with neighborhood invariance for universal domain adaptation," in *Proc. AAAI*, 2024, vol. 38, no. 4, pp. 3900–3908.
- [28] B. Fu, Z. Cao, M. Long, and J. Wang, "Learning to detect open classes for universal domain adaptation," in *Proc. ECCV*. Glasgow, U.K.: Springer, 2020, pp. 567–583.
- [29] W. Chang, Y. Shi, H. Tuan, and J. Wang, "Unified optimal transport framework for universal domain adaptation," in *Proc. NeurIPS*, vol. 35, 2022, pp. 29512–29524.
- [30] H. Zhang, M. Cisse, Y. N. Dauphin, and D. Lopez-Paz, "Mixup: Beyond empirical risk minimization," 2017, *arXiv:1710.09412*.
- [31] X. Peng, X. Zhu, and Y. Ma, "CL3D: Unsupervised domain adaptation for cross-LiDAR 3D detection," in *Proc. AAAI*, 2023, vol. 37, no. 2, pp. 2047–2055.
- [32] A. Sahoo, R. Panda, R. Feris, K. Saenko, and A. Das, "Select, label, and mix: Learning discriminative invariant feature representations for partial domain adaptation," in *Proc. IEEE/CVF Winter Conf. Appl. Comput. Vis. (WACV)*, Jan. 2023, pp. 4210–4219.
- [33] W. Tranheden, V. Olsson, J. Pinto, and L. Svensson, "DACS: Domain adaptation via cross-domain mixed sampling," in *Proc. IEEE/CVF Winter Conf. Appl. Comput. Vis.*, Jan. 2021, pp. 1379–1389.
- [34] V. Olsson, W. Tranheden, J. Pinto, and L. Svensson, "ClassMix: Segmentation-based data augmentation for semi-supervised learning," in *Proc. IEEE/CVF Winter Conf. Appl. Comput. Vis.*, Jan. 2021, pp. 1369–1378.
- [35] V. Verma et al., "Manifold mixup: Better representations by interpolating hidden states," in *Proc. Int. Conf. Mach. Learn.*, 2019, pp. 6438–6447.
- [36] E. D. Cubuk, B. Zoph, D. Mané, V. Vasudevan, and Q. V. Le, "AutoAugment: Learning augmentation strategies from data," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 113–123.
- [37] D. Ho, E. Liang, X. Chen, I. Stoica, and P. Abbeel, "Population based augmentation: Efficient learning of augmentation policy schedules," in *Proc. Int. Conf. Mach. Learn.*, 2019, pp. 2731–2741.
- [38] E. D. Cubuk, B. Zoph, J. Shlens, and Q. V. Le, "RandAugment: Practical automated data augmentation with a reduced search space," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2020, pp. 702–703.
- [39] G. French, M. Mackiewicz, and M. Fisher, "Self-ensembling for visual domain adaptation," in *Proc. ICLR*, 2018, pp. 1–20.
- [40] L. Xiao et al., "Self-supervised domain adaptation with consistency training," in *Proc. 25th Int. Conf. Pattern Recognit. (ICPR)*, Jan. 2021, pp. 6874–6880.
- [41] M. Galar, A. Fernández, E. Barrenechea, H. Bustince, and F. Herrera, "An overview of ensemble methods for binary classifiers in multi-class problems: Experimental study on one-vs-one and one-vs-all schemes," *Pattern Recognit.*, vol. 44, no. 8, pp. 1761–1776, Aug. 2011.
- [42] H. Liu et al., "Action understanding based on a combination of one-versus-rest and one-versus-one multi-classification methods," in *Proc. 10th Int. Congr. Image Signal Process., Biomed. Eng. Informat. (CISP-BMEI)*, Oct. 2017, pp. 1–5.
- [43] J. Jang and C. Ouk Kim, "One-vs-rest network-based deep probability model for open set recognition," 2020, *arXiv:2004.08067*.
- [44] S. Padhy, Z. Nado, J. Ren, J. Liu, J. Snoek, and B. Lakshminarayanan, "Revisiting one-vs-all classifiers for predictive uncertainty and out-of-distribution detection in neural networks," 2020, *arXiv:2007.05134*.
- [45] C. Luo, C. Song, and Z. Zhang, "Learning to adapt across dual discrepancy for cross-domain person re-identification," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 2, pp. 1963–1980, Feb. 2023.
- [46] Y. Elkin and V. Kurlin, "A new near-linear time algorithm for k-nearest neighbor search using a compressed cover tree," in *Proc. Int. Conf. Mach. Learn.*, 2023, pp. 9267–9311.
- [47] L. Peterson, "K-nearest neighbor," *Scholarpedia*, vol. 4, no. 2, p. 1883, 2009.
- [48] N. Khodadadi et al., "BAOA: Binary arithmetic optimization algorithm with K-nearest neighbor classifier for feature selection," *IEEE Access*, vol. 11, pp. 94094–94115, 2023.
- [49] V. A. Gonzalez-Huitron, A. E. Rodriguez-mata, L. E. Amabilis-Sosa, R. Baray-Arana, I. Robledo-Vega, and G. Valencia-Palomo, "Jaccard distance as similarity measure for disparity map estimation," *IEEE Latin Amer. Trans.*, vol. 21, no. 5, pp. 690–698, May 2023.
- [50] C. Debes et al., "Hyperspectral and LiDAR data fusion: Outcome of the 2013 GRSS data fusion contest," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 2405–2418, Jun. 2014.
- [51] B. Le Saux, N. Yokoya, R. Hansch, and S. Prasad, "2018 IEEE GRSS data fusion contest: Multimodal land use classification [technical committeees]," *IEEE Geosci. Remote Sens. Mag.*, vol. 6, no. 1, pp. 52–54, Mar. 2018.
- [52] Y. Ganin et al., "Domain-adversarial training of neural networks," *J. Mach. Learn. Res.*, vol. 17, no. 1, pp. 2030–2096, May 2016.
- [53] M. Long, Z. Cao, J. Wang, and M. I. Jordan, "Conditional adversarial domain adaptation," in *Proc. NeurIPS*, vol. 31, 2018, pp. 1–11.
- [54] K. Saito, D. Kim, S. Sclaroff, and K. Saenko, "Universal domain adaptation through self supervision," in *Proc. Adv. Neural Inf. Process. Syst.*, 2020, pp. 16282–16292.
- [55] J. Yang, J. Liu, N. Xu, and J. Huang, "TVT: Transferable vision transformer for unsupervised domain adaptation," in *Proc. IEEE/CVF Winter Conf. Appl. Comput. Vis. (WACV)*, Jan. 2023, pp. 520–530.
- [56] Z. Xiao et al., "SPA: A graph spectral alignment perspective for domain adaptation," in *Proc. NeurIPS*, vol. 36, 2024, pp. 1–21.
- [57] G.-S. Xia et al., "AID: A benchmark data set for performance evaluation of aerial scene classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 7, pp. 3965–3981, Jul. 2017.
- [58] G. Cheng, J. Han, and X. Lu, "Remote sensing image scene classification: Benchmark and state of the art," *Proc. IEEE*, vol. 105, no. 10, pp. 1865–1883, Oct. 2017.
- [59] Y. Yang and S. Newsam, "Bag-of-visual-words and spatial extensions for land-use classification," in *Proc. 18th SIGSPATIAL Int. Conf. Adv. Geograph. Inf. Syst.*, Nov. 2010, pp. 270–279.
- [60] Y. Liang, S. Cao, J. Zheng, X. Zhang, J. Huang, and H. Fu, "Low saturation confidence distribution-based test-time adaptation for cross-domain remote sensing image classification," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 139, May 2025, Art. no. 104463.
- [61] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.
- [62] Y. Zhang, W. Li, M. Zhang, Y. Qu, R. Tao, and H. Qi, "Topological structure and semantic information transfer network for cross-scene hyperspectral image classification," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 6, pp. 2817–2830, Jun. 2023.
- [63] M. Long, H. Zhu, J. Wang, and M. I. Jordan, "Deep transfer learning with joint adaptation networks," in *Proc. Int. Conf. Mach. Learn.*, 2017, pp. 2208–2217.
- [64] L. Van der Maaten and G. Hinton, "Visualizing data using t-SNE," *J. Mach. Learn. Res.*, vol. 9, no. 11, pp. 2579–2605, 2008.