HyUniDA: Breaking Label Set Constraints for Universal Domain Adaptation in Cross-Scene Hyperspectral Image Classification

Qingmei Li, Yibin Wen, Juepeng Zheng, *Member, IEEE*, Yuxiang Zhang,
and Haohuan Fu, *Senior Member, IEEE*

Abstract— Although enormous domain adaptation (DA) approaches have been proposed for cross-scene hyperspectral image (HSI) classification, the majority of DA methods strongly depend on much prior knowledge of the association among the label sets of source and target domains (encompassing closed set, partial, and open set DA), thereby significantly hindering their applications. Realistic application scenarios often require knowledge transfer between domains without restrictions on the label space, which is called universal domain adaptation (UniDA). In this article, we propose HyUniDA, which is the first attempt to address UniDA scenario from HSIs. HyUniDA contains two major parts: shared semantic pairing (SSP) and domain similarity score (DSS). We group both source and target domains to form discriminative clusters. The SSP identifies pairs of clusters that have coincident semantic features as the common classes. By examining the consistency level of samples across source and target domains, DSS can estimate the quantity of target clusters and generate distinct clusters without prior knowledge. Meanwhile, we apply the contrastive domain discrepancy to alleviate the offset of samples distribution, with a representative regularizer to assist in distinguishing target-domain clusters. We evaluate our proposed method on three transfer learning tasks for six typical HSI datasets; it turns out that our proposed method yields 3.83%–37.57% improvements compared to other state-of-the-art (SOTA) DA methods.

Index Terms— Cross scene, domain alignment, hyperspectral image (HSI) classification, shared semantic pairing (SSP), universal domain adaptation (UniDA).

Qingmei Li is with the College of Engineering, Peking University, Beijing 100871, China (e-mail: qmli@stu.pku.edu.cn).

Yibin Wen and Juepeng Zheng are with the School of Artificial Intelligence, Sun Yat-sen University, Zhuhai 510275, China (e-mail: wenyb5@mail2.sysu.edu.cn; zhengjp8@mail.sysu.edu.cn).

Yuxiang Zhang is with the School of Information and Electronics, Beijing Institute of Technology, Beijing 100081, China (e-mail: zyx829625@163.com).

Haohuan Fu is with the Tsinghua Shenzhen International Graduate School, Tsinghua University, Shenzhen 518071, China, also with the National Supercomputing Center in Shenzhen, Shenzhen 518055, China, also with the Ministry of Education Key Laboratory for Earth System Modeling and the Department of Earth System Science, Tsinghua University, Beijing 100084, China, and also with Tsinghua University (Department of Earth System Science)–Xi'an Institute of Surveying and Mapping Joint Research Center for Next-Generation Smart Mapping, Beijing 100084, China (e-mail: haohuan@tsinghua.edu.cn).

Digital Object Identifier 10.1109/TGRS.2024.3400959

I. INTRODUCTION

HYPERSPECTRAL images (HSIs) possess vast amounts of spatial texture and spectral features. They can obtain hundreds of spectral bands spanning from ultraviolet to infrared, and exhibit a remarkable capacity for information detection. HSI classification has been a research hotspot for remote sensing image processing [1], with extensive applications in geological exploration [2], environmental monitoring [3], and smart agriculture [4], [5]. However, the spectral information in HSI mostly does not display a linear relation with the materials of the land surface.

Deep learning has been acknowledged as a capable feature extraction tool, which can tackle nonlinear problems efficiently and show excellent performance in tasks such as image classification [6], [7], [8], [9], object detection [10], [11], [12], [13], [14], and image segmentation [15], [16], [17], [18]. The rapid advancement of these algorithms has led to the introduction of some deep learning-based methods in HSI classification, which have been proven to be superior on HSI datasets [19], [20], [21]. In HSI classification, reliance on abundant labeled datasets is essential, notably for training deep neural networks with numerous parameters [22]. Furthermore, traditional machine learning-based HSI classification assumes the *independence* and *identical distribution* (IID) principle, requiring the training data (source domain) and testing data (target domain) to be independent and identically distributed [23], [24]. Maintaining IID is crucial to prevent performance degradation caused by domain distribution gaps [25]. The time-consuming and laborious for labeling samples raises a major barrier to HSI classification, and the spectral information of HSI varies with the season and weather conditions of data acquisition, resulting in a shift between the source and target domains. Recent studies have demonstrated the feasibility of classification with limited sample sizes. For instance, Zhu et al. [26] proposed a spectral–spatial-dependent global learning framework, addressing the limitations posed by a scarcity of samples. Numerous convolutional neural network (CNN)-based classification methods face challenges in generalizing effectively across cross-domain tasks, particularly where the training and testing samples originate from different data distributions [27], [28], [29], [30].

Transfer learning is particularly useful to address this issue, where domain adaptation (DA) can transfer knowledge via

a completely annotated source to an unknown target domain by narrowing the distribution gap among source and target domains [31], [32], [33], [34]. Unsupervised DA allows to leverage labeled data from other, more easily accessible domains, to improve performance in the target domain with unlabeled data. Yan et al. [35] take the changes of class prior distributions into account and add class-specific auxiliary weight bias to maximum mean discrepancy (MMD). Besides that, adversarial learning methods have also been applied in DA to train a model that cannot be distinguished by the domain classifier to learn domain invariant features. Pei et al. [36] propose the multi-adversarial domain adaptation (MADA) method to capture multimode structures based on multiple domain discriminators. Long et al. [37] propose conditional domain adversarial networks (CDANs) to condition the adversarial adaptation models on discriminative information. Cross-scene HSI classification methods using DA have been widely applied [38], [39]. In the HSI community, Deng et al. [40] propose multikernel learning with active learning (MKL-AL) to retrain a multikernel classifier with a small number of labeled samples by active learning. Wang et al. [41] propose domain adaptation broad learning (DABL), which uses MMD in mapped features and adds manifold regularization to the input and output layer. Ma et al. [24] introduce three modules for domain alignment, task distribution, and DA to minimize the domain discrepancy and transfer ability of the source model to the target domain.

Most existing HSI classification methods adhere to a strict and ideal assumption. It is required that identical label sets should be shared between the source domain and target domain, that is, all categories of ground objects in the target domain must appear in the training set. This is the closed environment, and the DA in this scenario is named the closed set DA, as depicted in Fig. 1(a) [42]. Due to the diversity of land covers, this assumption may not be valid in some cases, and categories absent in the training set may emerge unexpectedly in the test set. Typical closed-set DA methods employ a linear classification layer or softmax function, which usually yields the maximum probability for an unknown class according to the known class. As illustrated in Fig. 1(b), the open set setting involves a scenario where the classes recognized in the source domain form a segment of the classes observed by the target domain. The DA in an open set environment is designed to effectively and robustly handle the unknown class of the target domain (i.e., target private class) [43], [44], [45]. Yue et al. [43] propose a spectral–spatial latent reconstruction framework to achieve robust unknown detection. Xie et al. [44] present feature consistency-based prototype network (FCPN) with a prototype-guided open set module to identify the outliers. What is more, the uncommon categories may appear in the source domain (i.e., source private class). As illustrated in Fig. 1(c), a notable observation is that the source domain contains a set of labels for the target domain, representing the partial environment [46], [47]. DA for these scenarios is termed the partial DA.

The aforementioned DA settings might face challenges in terms of rationality and realism when dealing with an unsupervised target domain. Furthermore, universal DA (UniDA)
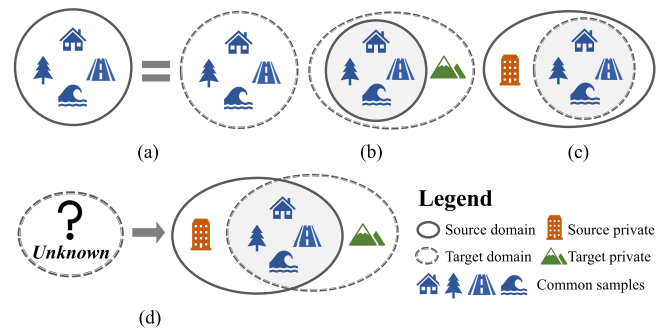

Fig. 1. Illustration of different DA settings. (a) Closed set DA. The classes in both the source and target domains exactly match. (b) Open set DA. The label set of the source domain is a subset of the target domain, signifying the existence of *unknown* class in the target domain. (c) Partial DA. The source domain contains a set of labels for the target domain, i.e., the *unknown* class may appear in the source domain. (d) Universal DA. Both source and target domains have individual private classes, but without knowing any prior information.

generalizes the above three DA settings [48], [49], [50]. UniDA defines a scenario where the source and target domains usually share some labels, but at the same time, each has a private set of labels that the other does not have, which is not restricted to any prior knowledge, as shown in Fig. 1(d). You et al. [48] propose the idea of UniDA and design universal adaptation network (UAN) to discover the common and private label sets by exploiting both domain similarity and prediction uncertainty for each sample. Saito et al. [51] present domain adaptive neighborhood clustering via entropy optimization (DANCE) to learn the structure of the target domain with the neighborhood clustering method. Chang et al. [52] propose unified optimal transport framework for UniDA (UniOT) to find common classes without hand-tuned threshold and use global statistical information of the assignment matrix to distinguish common and private classes.

The scenario setting of UniDA is more practical and common in the field of remote sensing. For instance, researchers have a model trained on the labeled dataset of the Amazon rainforest and hope to apply it to a new study area, such as the Himalayas. The new region differs from the previous study area and may contain categories that have never appeared in training datasets, such as alpine meadows and bare rock. Meanwhile, the target domain lacks swamp and wetland areas that are unique to the Amazon. Another common task is land cover mapping. Metropolitan areas like New York, NY, USA, and Tokyo, Japan, always attract the attention of researchers, and their land cover-type maps are easily obtained. When we want to study a microcity, such as Sedona, DA can greatly reduce the cost of mapping. However, the land types in the two cities are very different, and the Red Rock is also privately owned by Sedona. Efficiently distinguishing the common samples from the private samples of both domains is a major challenge in transferring across label spaces that are out of alignment. UniDA can cope with land-type changes across different cities and environments and help models learn transfer between different areas, enabling them to identify and adapt to new feature types. Notably, Xu et al. [53] proposed a UniDA method for RGB image classification, model adaptation domain adaptation (MA) and source data generation–model adaptation (SDG-MA), which manually set
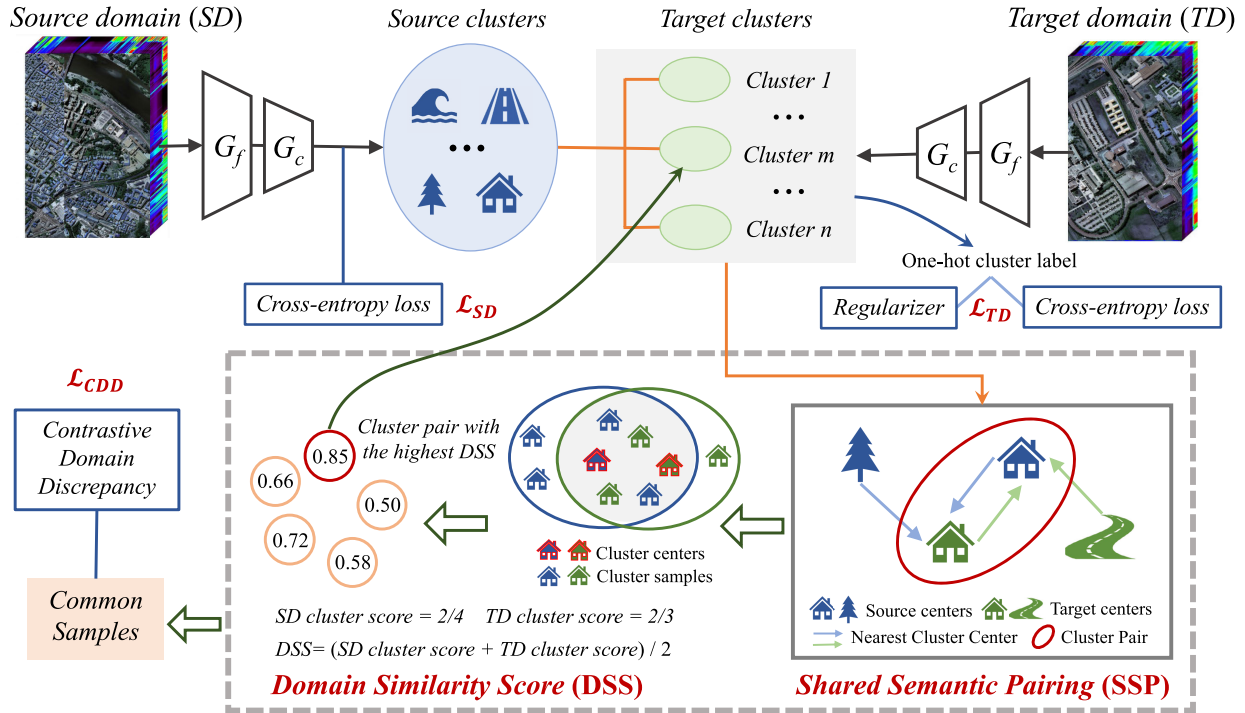
Fig. 2. Structure of our proposed HyUniDA, which involves two major parts: i.e., SSP and DSS. We group the instances from both the source and target domains into distinct clusters. Using the SSP to identify common classes among target clusters, if two clusters from both domains serve as the closest neighbors to each other, the instances belonging to the pairs of clusters are recognized as common ones that share the identical semantic annotations. In order to select the optimal target cluster, we establish a new metric DSS, which represents the degree of consistency among the cluster pairs. To the end, we apply the CDD to alleviate the offset of samples distribution, with a representative regularizer to assist in distinguishing target-domain clusters.

thresholds to detect common and unknown samples, limiting their extension to more practical scenarios. In the HSI community, while various methodologies have been advanced to address the challenges posed by DA scenarios [54], [55], seldom work is devoted to this important and practical but challenging situation (UniDA setting).

To address the aforementioned problems, we propose a UniDA framework based on the shared semantic pairing (SSP) and domain similarity score (DSS) for cross-scene HSI classification, namely, HyUniDA, to identify the class structure of the target domain. We cluster the source domain and target domain, respectively, to form discriminative clusters. Using SSP to explore semantic information for sample clusters, and when two cluster centers simultaneously act as centers closest to each other, the pair will be considered a common cluster. Next, we design the DSS to characterize the proportion of samples that achieve semantic agreement. By examining the consistency level of samples across source and target domains, DSS can estimate the number of target clusters and generate distinct clusters without prior knowledge. For those centers that are unable to locate the cluster peers, we use a prototypical regularizer to assist them in contacting attached centers. This work provides the following key contributions to the HSI field.

1) Aiming at the inconsistency of label space in cross-scene HSI classification, we propose the HyUniDA to solve the UniDA from the perspective of common class detection. As per our current awareness, this work stands out as the first initiative to bring the UniDA setting to the HSI field.

2) The proposed HyUniDA consists of SSP and DSS. SSP identifies cluster pairs with coincident semantic features as common classes, while DSS estimates the number of target clusters and generates distinct clusters without prior knowledge, enhancing the model adaptability in HSI.

3) The model incorporates contrastive domain discrepancy (CDD) to alleviate sample distribution offset and introduces a regularizer for effective target-domain cluster distinction, which enhances the robustness and efficacy of the proposed framework in addressing challenges associated with sample distribution shifts in HSI classification.

The subsequent sections are organized as follows. Section II delves into the details of the proposed HyUniDA methods, followed by the exposition of our datasets in Section III. The performance evaluation, along with comprehensive experiments and analyses, is undertaken in Section IV. Finally, we conclude the article in Section VI.

II. METHODOLOGY

Fig. 2 is employed to showcase an outline of the proposed HyUniDA. First, we cluster the target samples based on the K-means and get centers of the candidate clusters. We use the SSP to determine the clusters for that represent common classes in both domains within the obtained target clusters. When two clusters, each representing a distinct domain, become the closest counterparts to one another, the instances belonging to the pairs of clusters are recognized as common
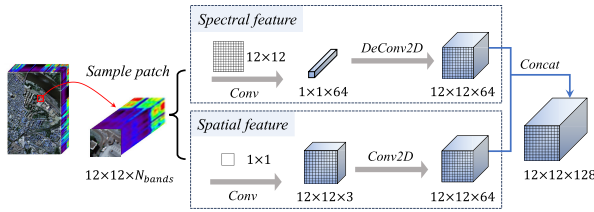
Fig. 3. Semantic encoder for the spatial and spectral construction processes. The 12×12 convolution kernel integrates spectral details, producing spectral embedding features f_{spe} ($1 \times 1 \times 64$). A subsequent deconvolution 2-D operation transforms the spectral feature into a $12 \times 12 \times 64$ feature map. Simultaneously, the spatial patch is condensed via a 1×1 convolution kernel. The spatial feature is mapped back to a $12 \times 12 \times 64$ feature map.

ones that share the identical semantic annotations. Then, the DSS is utilized to quantify the consistency among the paired clusters, and the cluster with the higher score indicates better consistency with the source-domain cluster. Lastly, we apply the CDD to alleviate the offset of samples distribution, with a representative regularizer to assist in distinguishing target-domain clusters.

A. Problem Setting

Given that $\mathcal{D}^s = \{(\mathbf{x}_i^s, \mathbf{y}_i^s)\}_{i=1}^{n_s}$ is the labeled source domain, where $\mathbf{x}_i^s$ is the data, $\mathbf{y}_i^s$ is the corresponding label, and n_s denotes the number of source samples. The target domain is unlabeled $\mathcal{D}^t = \{(\mathbf{x}_i^t)\}_{i=1}^{n_t}$. For $\mathbf{x}_i^s, \mathbf{x}_i^t \in \mathbb{R}^{H \times W \times B}$ and $\mathbf{y}_i^s \in \{0, 1, 2, \ldots, K_s\}^{H \times W}$, H and W correspond to the height and width of the HSI, respectively, B represents the spectral band size, and K_s represents the amount of ground-truth classes. The annotation sets for the source and target domains are referred to as $\mathcal{C}_s$ and $\mathcal{C}_t$, separately. The label set that encompasses classes owned by both domains is denoted as $\mathcal{C} = \mathcal{C}_s \cup \mathcal{C}_t$, and $\overline{\mathcal{C}}_s = \mathcal{C}_s \backslash \mathcal{C}$ stands for a set of annotations specific to the source domain. The private label set assigned to the targeted domain is denoted as $\overline{\mathcal{C}}_t = \mathcal{C}_t \backslash \mathcal{C}$, which is regarded as *unknown* class. Our goal is to identify the target instances into any given label from the $\mathcal{C}_s$ or the *unknown* class.

B. Shared Semantic Pairing

1) Semantic Encoder: In light of the robust spatial recognition and diverse multiband spectral details present in HSI data, the semantic encoder initiates a workflow for creating spatial and spectral dimensions, as shown in Fig. 3. The integration of spectral information contained in the patch is accomplished by a 12×12 convolution kernel during the spectral generation process, yielding spectral embedding features f_{spe} with dimensions $1 \times 1 \times 64$. Utilizing the deconvolution 2-D operation, the spectral feature ($\mathcal{S}_{pe}$) undergoes a transformation, resulting in a feature map sized $12 \times 12 \times 64$. The source domain contributes the spatial patch of $12 \times 12 \times d$ in the HSI dataset, which is then subjected to semantic encoding by the generator. Employing the 1×1 convolution kernel leads to the dimension of the spatial patch being condensed to $12 \times 12 \times 3$. Feature map f_{spa} represents spatial information. The spatial feature ($\mathcal{S}_{pa}$) is mapped back to a $12 \times 12 \times 64$ feature map.

We group the source- and target-domain samples into clusters, respectively. The center of jth source cluster, ζ_j^s,

can be written as

$$\zeta_j^s = \frac{1}{n_j^s} \sum_{\mathbf{x}_i^s \in \mathcal{D}_j^s} \frac{G_f(\mathbf{x}_i^s)}{\|G_f(\mathbf{x}_i^s)\|} \tag{1}$$

where $\mathcal{D}_j^s = \{\mathbf{x}_i^s\}_{i=1}^{n_j^s}$, feature extractor G_f maps the input images into vector representation, and ζ_j^s is computed by the weighted average of feature vectors across all samples in the cluster. For the purpose of clustering target samples, we employ K-means to categorize them into K groups, and the associated centers are denoted as $\{\zeta_1^t, \ldots, \zeta_K^t\}$.

Distinguishing between public and private samples is the main difficulty in UniDA. The construction of the SSP is devised to foster semantic-level associations among clusters that belong to identical classes. We seek the most proximate cluster center within the other domain for each cluster center. When the two clusters are simultaneously the closest to each other, they form the pair of common cluster.

By means of K-means clustering, we determine the original prototypes corresponding to target samples, $\{\zeta_{1(0)}^t, \ldots, \zeta_{K(0)}^t\}$. Herein, we present a detailed evaluation mechanism for prototypes during the training process. At every iteration, the determination of localized prototypes within a batch is explicated as follows:

$$\overline{\zeta}_{k(\mathbf{I})}^t = \frac{1}{|\overline{\mathcal{D}}_{k(\mathbf{I})}^t|} \sum_{\mathbf{x}_i^t \in \overline{\mathcal{D}}_{k(\mathbf{I})}^t} \frac{G_f(\mathbf{x}_i^t)}{\|G_f(\mathbf{x}_i^t)\|} \tag{2}$$

where $\overline{\zeta}_{k(\mathbf{I})}^t$ is calculated in a similar way to ζ_j^s above, $\mathbf{I}$ represents the running iteration, and $\overline{\mathcal{D}}_{k(\mathbf{I})}^t$ indicates target samples characterized by the cluster label k during the iteration $\mathbf{I}$. The global prototype undergoes the following update:

$$\zeta_{k(\mathbf{I})}^t = \delta_{\mathbf{I}} \zeta_{k(\mathbf{I}-1)}^t + (1 - \delta_{\mathbf{I}}) \overline{\zeta}_{k(\mathbf{I})}^t. \tag{3}$$

The update results from a weighted fusion of the previous global instance $\zeta_{k(\mathbf{I}-1)}^t$ and the average prototype $\overline{\zeta}_{k(\mathbf{I})}^t$ at the current. $\delta_{\mathbf{I}}$ is the cosine similarity among the local instances and global instances, which can be expressed as follows:

$$\delta_{\mathbf{I}} = \frac{\left\langle \zeta_{k(\mathbf{I}-1)}^t, \overline{\zeta}_{k(\mathbf{I})}^t \right\rangle}{\left\| \zeta_{k(\mathbf{I}-1)}^t \right\| \left\| \overline{\zeta}_{k(\mathbf{I})}^t \right\|}. \tag{4}$$

Through an adaptive mechanism driven by the agreement between global and local instances, the prototype bank undergoes updates in a manner that maximizes efficiency.

C. Optimization of Target Clusters

While the SSP excels at recognizing common classes, gauging the quantity of target clusters proves challenging in the absence of information regarding the class distribution in the target domain. Some studies have proposed clustering evaluation criteria to measure the number of clusters [56]. Unfortunately, these methods are tailored to a single domain and do not directly account for cross-scene knowledge. As a solution, we suggest using DSS to identify the number of target clusters based on the consistency across samples.

For a pair of clusters $\{p_i^s\}_{i=1}^m$ and $\{p_i^s\}_{i=1}^n$, whose centers are ζ_j^s and ζ_k^t, respectively, for each sample of the source

Algorithm 1 Unified UniDA Framework for HSI

Input: Source domain $\mathcal{D}^s = \left\{\left(\mathbf{x}_i^s, \mathbf{y}_i^s\right)\right\}_{i=1}^{n_s}$, target domain $\mathcal{D}^t = \left\{\left(\mathbf{x}_i^t\right)\right\}_{i=1}^{n_t}$, and trade-off parameter λ, ω.

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: **while** not converged **do**

3:     Train source domain and calculate the ($\mathcal{L}_{SD}$) according to Equation (8);

4:     Cluster pairing, the cluster centers for source and target are calculated according to Equation (1) and (2) respectively.

5:     Cross-domain alignment on identified common samples with CDD, according to Equation (10)–(13);

6:     Train target domain and calculate the ($\mathcal{L}_{TD}$) according to Equation (14);

7:     **if** Iteration > 1 **then**

8:         Update the target centers according to Equation (3) and (4);

9:     **end if**

10:    Update the feature extractor G_f and classifier G_c.

11: **end while**

12: Obtain the predicted label $\hat{\mathbf{y}}^t$ for target domain.

---

domain, the distance between it and the target cluster centers $\{\zeta_1^t, \ldots, \zeta_K^t\}$ is calculated

$$d_{i,k}^s = \frac{\left\langle p_i^s, \zeta_k^t \right\rangle}{\left\| p_i^s \right\| \left\| \zeta_k^t \right\|}, \quad k \in \{1, \ldots, K\}. \tag{5}$$

The score of sample clustering, $\text{SCS}_{(j,k)}^s$, is the proportion of those samples in the clustered pair formed by the SSP that appear in both domains. It can be expressed as follows:

$$\text{SCS}_{(j,k)}^s = \frac{\sum_{i=1}^m \mathbf{1}\left\{\arg\max_k\left(d_{i,k}^s\right) = k\right\}}{m} \tag{6}$$

when the condition is true, the binary indicator $\mathbf{1}\{\cdot\}$ assumes a value of 1; otherwise, it takes on a value of 0. Similarly, the score of sample clustering for the target domain $\text{SCS}_{(j,k)}^t$ can be obtained. The DSS is the average of the two scores

$$\text{DSS}_{(j,k)} = \frac{\text{SCS}_{(j,k)}^s + \text{SCS}_{(c,k)}^t}{2}. \tag{7}$$

Finally, we calculate the average DSS across all corresponding cluster pairs. The optimal number of target clusters is deduced by conducting clustering experiments with different values of K. We perform the DSS calculation over a range of K values at regular intervals and opt for the one demonstrating the utmost score to guide the ensuing clustering process.

We aim to ascertain the value of K in one search, but the vast domain gap at the initial training phase makes this unfeasible. The DSS initially rises and then falls with the increase of K value. To enhance search efficiency, the search is halted when the score decreases several times consecutively, and a fixed value of K is retained at certain rounds.

D. Domain Alignment

Cross-scene HSI classification faces a substantial challenge attributed to spectral shift. Spectral reflectance of the identical land cover type can often differ in the source and target HSIs. The feature generator G_f can be applied to obtain the discriminant features of the source domain, after which the classifier G_c divides the samples into different categories. Throughout the HyUniDA training process, the primary emphasis lies in minimizing the loss associated with label predictions on labeled HSI cubes originating from the source domain. This strategic optimization aims to refine the parameters of both G_f and G_c for the purpose of decreasing experiential loss of the source HSI. The classification loss for source domain ($\mathcal{L}_{SD}$) can be formulated as follows:

$$\begin{aligned}
\mathcal{L}_{SD} &= \frac{1}{n_s} \sum_{i=1}^{n_s} \mathcal{L}_{CE}\left(G_c\left(G_f\left(x_i^s\right)\right), y_i^s\right) \\
&= -\frac{1}{n_s} \sum_{i=1}^{n_s} \sum_{c=1}^{|C_s|} \hat{y}_{i,c}^s \log\left(\pi\left(G_c\left(G_f\left(x_i^s\right)\right)\right)\right)
\end{aligned} \tag{8}$$

where $\mathcal{L}_{CE}$ represents the standard cross-entropy loss, $\hat{y}_{i,c}^s$ represents the one-hot encoding correspond to the source label, and π signifies the Softmax function.

After the SSP has been completed, the identified common samples are gathered into clusters and we need to align these samples using a class-aware approach. MMD [57] delineates the disparity between the source and target domains in terms of the average embeddings distributed within the reproducing kernel Hilbert space (RKHS). The formulation is given by $\mathcal{D}_{\mathcal{H}}(P, Q) \triangleq \sup_{f \sim \mathcal{H}} \left(\mathbb{E}_{X^s}[f(X^s)] - \mathbb{E}_{X^t}[f(X^t)]\right)_{\mathcal{H}}$, where $\mathcal{H}$ is a class of functions, and $\{\mathbf{x}_i^s\}$ and $\{\mathbf{x}_i^t\}$ are sampled from the marginal distributions $P(X^s)$ and $Q(X^t)$, respectively. The MMD distance can be denoted as

$$\mathcal{D}^{\text{mmd}} = \left\| \frac{1}{n_s} \sum_{i=1}^{n_s} \phi\left(\mathbf{x}_i^s\right) - \frac{1}{n_t} \sum_{j=1}^{n_t} \phi\left(\mathbf{x}_j^t\right) \right\|_{\mathcal{H}}^2 \tag{9}$$

where the model is parameterized by ϕ. Despite MMD has worked well in measuring the marginal distribution difference between two domains, it focuses on domain-level distinctions, without consideration for the sample classes. It is unable to distinguish if samples of two domains exhibit alignment with the corresponding class labels.

CDD incorporates both the interclass and intraclass discrepancies based on MMD to realize class-aware alignment [58], [59]. CDD employs the maximum value of squared distance between the kernel means of two conditional distributions in RKHS and extends MMD by gauging the variance between $P(\phi(X^s)|Y^s)$ and $Q(\phi(X^t)|Y^t)$, that is, $\mathcal{D}_{\mathcal{H}}(P, Q) \triangleq \sup_{f \sim \mathcal{H}}(\mathbb{E}_{X^s}[f(\phi(X^s)|Y^s)] - \mathbb{E}_{X^t}[f(\phi(X^t)|Y^t)])_{\mathcal{H}}$. The empirical estimate for squared $\mathcal{D}_{\mathcal{H}}(P, Q)$ at the lth layer of CNN is given as follows ($\hat{y}_1^t, \hat{y}_2^t, \ldots, \hat{y}_{n_t}^t$ is abbreviated as $\hat{y}_{1:n_t}^t$):

$$\widehat{\mathcal{D}}^{c_1 c_2}\left(\hat{y}_{1:n_t}^t, \phi\right) = \Psi^{c_1 c_1}(s, s) + \Psi^{c_2 c_2}(t, t) - 2\Psi^{c_1 c_2}(s, t) \tag{10}$$

where

$$\Psi^{c_1 c_2}(s,t) = \sum_{i=1}^{n_s} \sum_{j=1}^{n_t} \frac{\mu_{c_1 c_2}(y_i^s, \hat{y}_j^t) k_l(\phi(x_i^s), \phi(x_j^t))}{\sum_{i=1}^{n_s} \sum_{j=1}^{n_t} \mu_{c_1 c_2}(y_i^s, \hat{y}_j^t)}$$

$$\mu_{c_1 c_2}(y_1, y_2) = \begin{cases} 1, & \text{if } y_1 = c_1, y_2 = c_2 \\ 0, & \text{otherwise.} \end{cases} \quad (11)$$

k_l represents the kernel applied to the lth layer in CNN; c_1 and c_2 are the class labels of source and target HSIs, respectively. When $c_1 = c_2 = c$, it captures domain discrepancy intraclass and strives to minimize the intraclass difference, thereby enhancing the compactness of sample features within the class. Whereas if $c_1 \neq c_2$, it represents domain discrepancy interclass and aims to maximize the difference between features of various classes, consequently enhancing their separation toward the decision-making boundary.

As per the quoted definition, CDD is quantified by the difference between discrepancy within classes ($\hat{\mathcal{D}}_l^{\text{intra}}$) and discrepancy among different classes ($\hat{\mathcal{D}}_l^{\text{inter}}$)

$$\hat{\mathcal{D}}_l^{\text{cdd}} = \hat{\mathcal{D}}_l^{\text{intra}} - \hat{\mathcal{D}}_l^{\text{inter}} = \frac{1}{|\mathcal{C}_s|} \sum_{c=1}^{|\mathcal{C}_s|} \hat{\mathcal{D}}^{cc}(\hat{y}_{1:n_t}^t, \phi)$$

$$- \frac{1}{|\mathcal{C}_s|(|\mathcal{C}_s| - 1)} \sum_{c=1}^{|\mathcal{C}_s|} \sum_{\substack{c'=1, \\ c' \neq c}}^{|\mathcal{C}_s|} \hat{\mathcal{D}}^{cc'}(\hat{y}_{1:n_t}^t, \phi). \quad (12)$$

Within a deep CNN, the objective is to reduce CDD across multiple fully-connected layers; specifically, the task involves minimizing

$$\mathcal{L}_{\text{CDD}} = \sum_{l=1}^{L} \hat{\mathcal{D}}_l^{\text{cdd}}. \quad (13)$$

CDD is a key component of HyUniDA, enhancing the model's ability to generalize across different datasets and scenarios. CDD minimizes intraclass differences and maximizes interclass margins, which is effective in HSI applications with spectral variations. By addressing class-level domain differences uniquely and optimizing them in a contrasting manner, HyUniDA attains enhanced domain alignment.

As with the source domain, we calculate the classification loss $\mathcal{L}_{\text{cls}}^T$ for the target domain by utilizing standard cross-entropy loss, thereby improving the discriminability of the model. In addition, we apply an entropy regularization loss $\mathcal{L}_{\text{reg}}^T$ upon the target scene to augment the distinctness of the target instances.

There is a prototype set $\Omega = [\zeta_1^t, \zeta_2^t, \ldots, \zeta_K^t]$, which contains L2-normalized centers of each target cluster, with ongoing iterative adjustments applied to the prototypes throughout the training process. The loss for target domain $\mathcal{L}_{\text{TD}}$ consists of the classification loss $\mathcal{L}_{\text{cls}}^T$ and the prototypical regularizer $\mathcal{L}_{\text{reg}}^T$, and it can be expressed as

$$\mathcal{L}_{\text{TD}} = \mathcal{L}_{\text{cls}}^T + \mathcal{L}_{\text{reg}}^T$$

$$= \frac{1}{n_t} \sum_{i=1}^{n_t} \mathcal{L}_{\text{CE}}\left(G_c\left(G_f\left(x_i^t\right)\right), \hat{y}_i^t\right) - \sum_{i=1}^{n_t} \sum_{k=1}^{K} \hat{y}_{i,k}^t \log \hat{p}_{i,k}$$

$$= -\frac{1}{n_t} \sum_{i=1}^{n_t} \sum_{c=1}^{|C_t|} \hat{y}_{i,c}^t \log\left(\pi\left(G_c\left(G_f\left(x_i^t\right)\right)\right)\right)$$

$$- \sum_{i=1}^{n_t} \sum_{k=1}^{K} \hat{y}_{i,k}^t \log \frac{\exp\left(v_i^T \zeta_k^t / \tau\right)}{\sum_{k=1}^{K} \exp\left(v_i^T \zeta_k^t / \tau\right)} \quad (14)$$

where $\hat{y}_i^t$ is the one-hot cluster label, $\hat{p}_{i,k}$ is the predicted probability by the classifier in the target domain, vector v_i represents the normalized feature for the target domain, and the temperature parameter τ consistently configured to 0.1.

E. Unified UniDA Framework for HSI (HyUniDA)

In our quest for UniDA in HSI community, we introduce a unified domain adaptation (HyUniDA) framework that seamlessly integrates three crucial components: the cross-entropy loss ($\mathcal{L}_{\text{SD}}$) on source samples, the loss associated with the target domain ($\mathcal{L}_{\text{TD}}$), and the CDD loss ($\mathcal{L}_{\text{CDD}}$). The training procedure of HyUniDA is depicted in Algorithm 1.

Our objective is to optimize a range of loss functions, which address the inherent complexities of HSI data and the challenges posed by varying source and target domains. The $\mathcal{L}_{\text{SD}}$ ensures effective utilization of source samples, adapting them to the distinctive features of the target domain. Meanwhile, the $\mathcal{L}_{\text{TD}}$ facilitates the alignment of the model with the distinctive characteristics inherent to the target domain. The $\mathcal{L}_{\text{CDD}}$ leverages class-aware distribution analysis, allowing to quantify the divergence among source and target domains more comprehensively. The overall objective is delineated via the following formulation:

$$\mathcal{L} = \mathcal{L}_{\text{SD}} + \gamma \mathcal{L}_{\text{TD}} + \lambda \mathcal{L}_{\text{CDD}}. \quad (15)$$

To enable the cluster number to grow in the initial training phase and avoid the acquisition of more private samples postsaturation, we introduce incremental parameters γ. $\gamma = e^{-\omega \times (2/N)}$, where i and N symbolize the current and global iteration, respectively, with $\omega = 3.0$. The λ is consistently configured at 0.1 for all datasets. Through the strategic integration of these loss components, we aim to provide an effective solution for handling the inherent challenges associated with diverse HSI datasets and domain shifts.

III. DATASETS

Experiments using six publicly available cross-scene HSI datasets, namely, Houston2013 and Houston2018, Pavia Center and University, and HyRANK Dioni and Loukia, are implemented to test the proposed HyUniDA. These datasets exhibit both geographical diversity and spectral complexity.

1) *Houston:* This dataset encompasses scenes acquired through different sensors, denoted as Houston2013 [60] and Houston2018 [61], surveying the University of Houston campus and its surroundings across different years. Focusing on the overlapping region measuring 209×955 pixels in the Houston2013 and Houston2018 scenes, we explored the commonality within 48 spectral bands. Seven classes are identified across both scenes. In this article, we select the *Nonresidential buildings* as source private class, *Road* as the target private class,
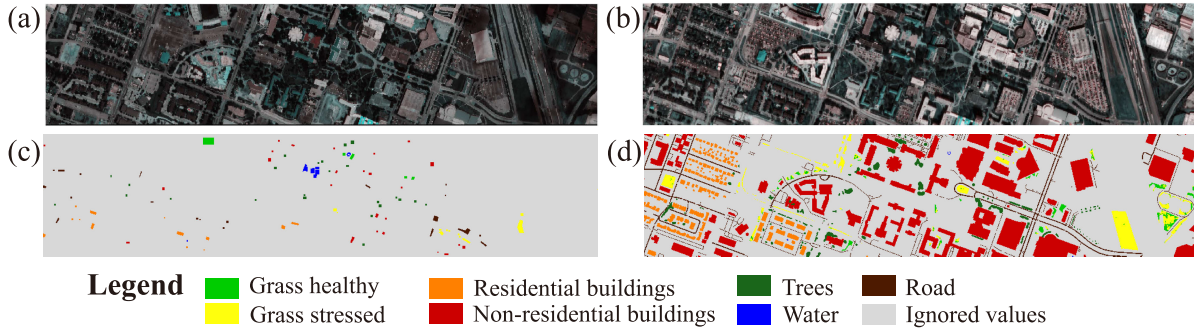
Fig. 4. Visualization for Houston task: (a) Houston2013 in pseudo-color; (b) Houston2018 in pseudo-color; (c) Houston2013 ground-truth map; and (d) Houston2018 ground-truth map.

TABLE I
NUMBER OF SAMPLES FOR HOUSTON DATASETS

Class		Number of Samples	
No.	Name	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

Class		Number of Samples	
No.	Name	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

and the other five classes (class 1–5) as common classes. Table I enumerates classes and their respective sample counts. Refer to Fig. 4 for visual depictions, presenting pseudo-color images alongside ground-truth maps.

2) *Pavia:* The dataset is composed of two components: Pavia Center (1096 × 715 pixels) and University (610 × 340 pixels). We abandoned the final spectral band in Pavia University original cube to obtain the Pavia dataset with consistent band, with a total of 102 bands. Both of them have the same seven classes. In this article, we select the *Meadow* as source private class, *Bare soil* as target private class, and the other five classes (class 1–5) as common classes. Table II enumerates classes and their respective sample counts. Refer to Fig. 5 for visual depictions, presenting pseudo-color images alongside ground-truth maps.

3) *HyRANK:* Derived from the Hyperion sensor, featuring a diverse range of 176 spectral bands [62], the HyRANK dataset presents two scenes marked with meticulous
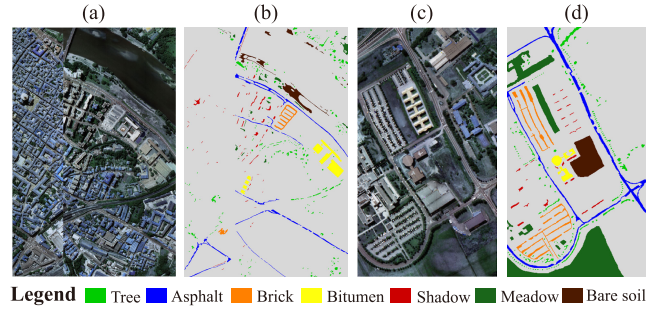

Fig. 5. Visualization for Pavia task: (a) Pavia University in pseudo-color; (b) Pavia Center in pseudo-color; (c) Pavia University ground-truth map; and (d) Pavia Center ground-truth map.

TABLE III
NUMBER OF SAMPLES FOR HYRANK DATASETS

Class		Number of Samples	
No.	Name	Dioni	Loukia
1	Dense Urban Fabric	1262	206
2	Mineral Extraction Sites	204	54
3	Non Irrigated Arable Land	614	426
4	Fruit Trees	150	79
5	Olive Groves	1768	1107
6	Coniferous Forest	361	422
7	Dense Sclerophyllous Vegetation	5035	2996
8	Sparse Sclerophyllous Vegetation	6374	2361
9	Sparsely Vegetated Areas	1754	399
10	Rocks and Sand	492	453
11	Water	1612	1393
12	Coastal Water	398	421
Total		20024	10317

labeling, Dioni and Loukia. There are 12 consistent classes for these two labeled scenes. In this article, we select *Sparsely Vegetated Areas* and *Rocks and Sand* as source private classes, *Water* and *Coastal Water* as target private classes, and the other eight classes (class 1–8) as common classes. Table III enumerates classes and their respective sample counts. Refer to Fig. 6 for visual depictions, presenting pseudo-color images alongside ground-truth maps.

IV. EXPERIMENTAL RESULTS

A. Setup

Our evaluation encompasses six HSI datasets to gauge the performance of the proposed method. Utilizing the
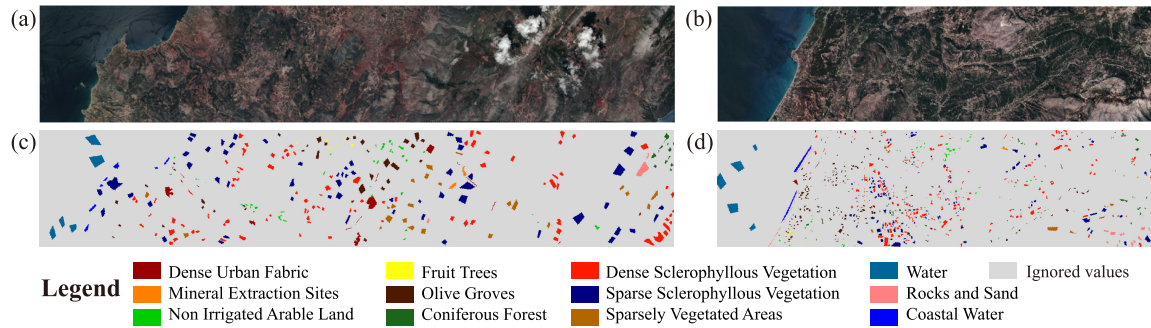
Fig. 6. Visualization for HyRANK task: (a) Dioni in pseudo-color; (b) Loukia in pseudo-color; (c) Dioni ground-truth map; and (d) Loukia ground-truth map.

PyTorch framework, we instantiate and execute HyUniDA, with the computations executed on the GeForce 1080 Ti GPU. We utilize the same architecture VGG16 as [38] for all our experiments in this article. The classifier includes two fully-connected layers adhering to the design of existing studies [48], [52]. The optimization process involves utilizing Nesterov momentum stochastic gradient descent (SGD) that takes a momentum at 0.9 and a weight decay parameter at 0.0005. The patch should be standardized to a size of 12×12, with each patch labeled according to the class of its geometrical center pixel, and the batch size is uniformly set to 128. A dynamic decay mechanism governs the learning rate, defined by $(1 + \alpha v)^{-\beta}$, with parameters $\alpha = 10$, $\beta = 0.75$, and $v = i/g$ signifying the linear transition in the training process, ranging from 0 to 1 in terms of the current iteration (i) to global iteration (g) ratio. The initial learning rate can be selected among {0.0001, 0.001, 0.01, 0.1}.

Evaluation Metric: For UniDA, the assessment of effectiveness requires a thoughtful consideration of the accuracy dynamics common and private categories. Besides reporting Class-specific/common/target private/overall classification accuracies, we also adopt the evaluation protocol H-score proposed by Fu et al. [63], which compensates for the per-class accuracy brought about by the previous neglect on open classes. Jointly assessing the accuracy of common class ($a_{\mathcal{C}}$) and target private class ($a_{\overline{\mathcal{C}}_t}$), the H-score offers a comprehensive performance measure. It can be computed as follows:

$$\text{H-score} = 2 \times \frac{a_{\mathcal{C}} \times a_{\overline{\mathcal{C}}_t}}{a_{\mathcal{C}} + a_{\overline{\mathcal{C}}_t}}. \tag{16}$$

Peak in this assessment hinges on the simultaneous increasing of both $a_{\mathcal{C}}$ and $a_{\overline{\mathcal{C}}_t}$, highlighting the proficiency in recognizing common and private samples.

B. Comparison Methods

To establish a comprehensive performance baseline, we incorporate a diverse set of state-of-the-art (SOTA) DA algorithms for comparison, highlighting the inclusion of two closed set DA methodologies, notably the domain-adversarial neural network (DANN) [64] and CDAN [37], one partial DA method (partial adversarial domain adaptation (PADA) [46]), one open set DA method (open set domain adaptation by backpropagation (OSBP) [65]), and four UniDA

methods (UAN [48], DANCE [51], UniOT [52], and UniDA for remote sensing image scene classification (MA) [53]). 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 highest probability is assigned to the target sample. When the probability of all known labels is below the threshold, we set it to *unknown*, i.e., target private. In addition, we include the Baseline model in our evaluation, which serves as a reference point for comparison. It should be emphasized that the Baseline model exclusively utilizes classification loss without incorporating any DA techniques.

Both source and target spectral data are row normalized using the L_2-norm during training and testing prior to network training. Tables IV–VI provide the common classes accuracy, target private accuracy, overall accuracy, and H-score in aforementioned methods for three target scenes, along with the classification accuracy for individual classes. We also list the average accuracy and standard deviation of the proposed method. The common class accuracy refers to the accuracy of those classes that appear in both the source and target domains, that is, the proportion of correctly identified common class samples to the total number of common class samples. Our HyUniDA exhibits superior performance across all benchmarks, leading in both overall accuracy and H-score. For closed-set DA methods, both comparison methods exhibit higher overall accuracy than Baseline, and most of them outperform the Baseline in terms of H-score, except for CDAN in Houston dataset (-0.1%) and DANN in HyRANK (-0.11%). DANN demonstrates high accuracy in specific categories, such as 76.66% for *Residential buildings* in Houston and 98.20% for *Shadow* in Pavia. OSBP scores highest for the accuracy of private samples within the target domain among the Pavia dataset. As for partial and open set DA, the algorithm may perform worse H-score than the naive Baseline, such as PADA in Houston dataset (-0.44%) and OSBP in HyRANK dataset (-8.36%), demonstrating a substantial negative transfer effect (TE) under the UniDA setting. While UniDA approaches yield superior results, it is important to note that negative transfer was still present. A surprising finding is that, as a recent SOTA method, UniOT demonstrates inferior performance in the UniDA scenario of HSI cross-domain classification, with H-score testing 3.08% lower than the Baseline in the

TABLE IV

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

Class	Methods									
	Baseline	DANN [64]	CDAN [37]	PADA [46]	OSBP [65]	UAN [48]	DANCE [51]	UniOT [52]	MA [53]	HyUniDA (Ours)
Grass healthy	2.07	10.50	9.31	5.99	45.31	29.34	**83.74**	54.84	82.93	73.44 ± 4.94
Grass stressed	43.56	52.82	47.71	43.86	36.29	**62.97**	52.99	26.55	41.35	50.62 ± 2.28
Trees	18.55	22.70	22.89	23.97	35.00	16.70	23.97	43.09	51.70	**76.55 ± 3.15**
Water	90.91	90.91	90.91	90.91	90.91	77.27	95.45	**100.00**	100.00	71.21 ± 9.46
Residential buildings	74.23	**76.66**	75.07	75.93	43.20	50.46	63.12	47.47	59.73	34.95 ± 2.78
Common classes	46.32	51.97	49.56	48.47	39.55	46.27	54.13	40.29	**54.18**	51.96 ± 2.65
Target private (Road)	55.18	52.98	50.98	51.47	76.64	71.39	58.65	74.27	57.38	**80.72 ± 1.84**
Overall accuracy	49.04	52.28	50.00	49.39	50.93	53.98	55.52	50.72	55.16	**60.79 ± 2.27**
H-score	50.36	52.47	50.26	49.92	52.17	56.15	56.30	52.24	55.73	**63.22 ± 2.38**

TABLE V

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

Class	Methods									
	Baseline	DANN [64]	CDAN [37]	PADA [46]	OSBP [65]	UAN [48]	DANCE [51]	UniOT [52]	MA [53]	HyUniDA (Ours)
Tree	41.16	68.18	60.38	59.79	30.16	54.86	70.40	69.61	62.37	**77.67 ± 4.92**
Asphalt	51.30	59.46	57.83	52.74	50.35	43.76	79.73	20.99	45.42	**93.84 ± 5.76**
Brick	11.62	7.03	5.13	6.57	**48.32**	25.94	29.44	2.04	39.73	7.24 ± 7.35
Bitumen	0.00	0.15	0.08	0.45	0.00	0.23	6.77	4.51	0.08	**20.53 ± 2.23**
Shadow	96.73	98.20	96.52	94.19	52.30	67.16	96.52	94.93	90.81	**98.73 ± 1.38**
Common classes	38.37	46.14	43.37	41.32	42.37	39.46	60.89	29.12	46.29	**64.37 ± 3.41**
Target private (Bare soil)	16.62	45.16	26.39	40.66	**72.22**	59.14	53.47	58.12	56.25	59.04 ± 2.17
Overall accuracy	33.09	45.90	39.24	41.16	49.63	44.24	59.09	36.17	48.72	**63.07 ± 3.01**
H-score	23.20	45.64	32.81	40.99	53.40	47.33	56.94	38.80	50.79	**61.59 ± 2.59**

TABLE VI

CLASSIFICATION ACCURACY METRICS, INCLUDING CLASS-SPECIFIC, COMMON, TARGET PRIVATE, OVERALL ACCURACY, AND H-SCORE (%), FOR VARIOUS SOTA DA METHODS EVALUATED FROM HyRANK DIONI TO THE TARGET SCENARIO LOUKIA

Class	Methods									
	Baseline	DANN [64]	CDAN [37]	PADA [46]	OSBP [65]	UAN [48]	DANCE [51]	UniOT [52]	MA [53]	HyUniDA (Ours)
Dense Urban Fabric	0.00	0.00	0.97	0.00	0.00	13.59	0.00	0.00	2.43	**38.54 ± 2.32**
Mineral Extraction Sites	9.26	42.59	0.00	12.96	1.85	42.59	**59.26**	0.00	0.00	32.59 ± 3.84
Non Irrigated Arable Land	14.55	3.99	27.00	10.33	8.45	10.56	7.75	**41.78**	15.73	31.55 ± 1.79
Fruit Trees	3.80	17.22	12.67	10.13	27.85	8.86	5.06	7.59	**37.97**	9.87 ± 6.96
Olive Groves	0.00	0.09	6.59	0.09	0.00	10.48	0.00	**74.89**	0.09	13.91 ± 0.51
Coniferous Forest	36.49	24.17	39.34	32.23	**48.82**	23.46	45.50	35.78	29.38	9.91 ± 7.32
Dense Sclerophyllous Vegetation	39.82	55.07	29.71	51.87	41.05	50.43	**63.12**	19.83	38.79	58.48 ± 3.61
Sparse Sclerophyllous Vegetation	21.22	0.72	46.21	20.84	0.00	53.24	39.34	0.13	21.01	**61.78 ± 6.90**
Common classes	25.07	23.84	30.68	29.30	19.54	40.33	40.27	23.02	24.64	**47.65 ± 1.87**
Target private (Water & Coastal water)	81.70	96.86	70.73	98.68	64.66	99.61	91.79	75.58	85.94	**99.78 ± 0.49**
Overall accuracy	35.92	37.83	38.35	42.60	28.19	51.70	50.14	33.09	36.39	**57.64 ± 1.51**
H-Score	38.37	38.26	42.79	45.19	30.01	57.42	55.98	35.29	38.30	**64.50 ± 1.70**

HyRANK dataset, and overall performance is even worse than that of closed set DA. The MA model designed for remote sensing RGB images also performed poorly, with an H-score of 38.30%. DANCE emerges as the most efficient of the three UniDA comparison methods we selected, with the highest accuracy of 54.13% for common classes in the Houston dataset. However, our proposed HyUniDA achieves the highest overall accuracy of 60.79%, 63.07%, and 57.64% and the highest H-score of 63.22%, 61.59%, and 64.50% for Houston, Pavia, and HyRANK, respectively, with notable gains (12.86%–38.39%) compared to the Baseline. Furthermore, compared with other SOTA methods, the HyUniDA attains 6.92%–13.30%, 4.65%–28.78%, and 7.08%–34.49% improvement with respect to H-score for these three HSI target scenarios.

Figs. 7–9 show the classification maps of each comparison algorithm for target scenarios Houston2018, Pavia University, and HyRANK Loukia, respectively. Within the depicted maps, pixels with labeling reflect the predicted class of ground objects, while the unlabeled ones indicate the background or ignored class. The black rectangles circle some key areas to demonstrate the superiority of the HyUniDA. Evidently, our proposed HyUniDA algorithm exhibits decreased noise and heightened precision in specific regions on classification maps. For instance, the third class (*Trees*) in the Houston2018, both the *Tree* and *Asphalt* classes in the Pavia University, and the *Dense Urban Fabric* class in the HyRANK Loukia, HyUniDA shows remarkable noise reduction and improved accuracy, outperforming other SOTA comparison methods. The proposed HyUniDA stands out for its precise capture of private samples within the target domain. It demonstrates the capability to correctly classify a greater number of samples, thereby enhancing the overall consistency between the generated classification maps and the ground-truth maps. This

Legend ■ Grass healthy ■ Grass stressed ■ Residential buildings ■ Trees ■ Water ■ Ignored values
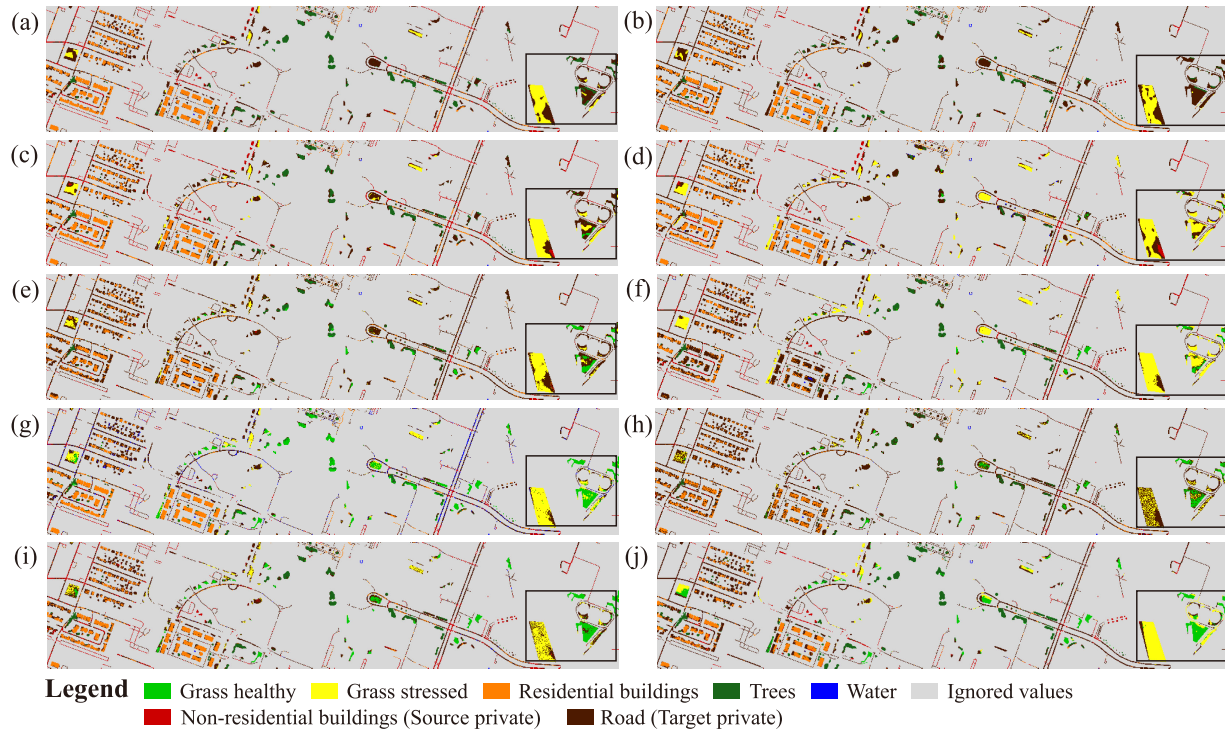■ Non-residential buildings (Source private) ■ Road (Target private)

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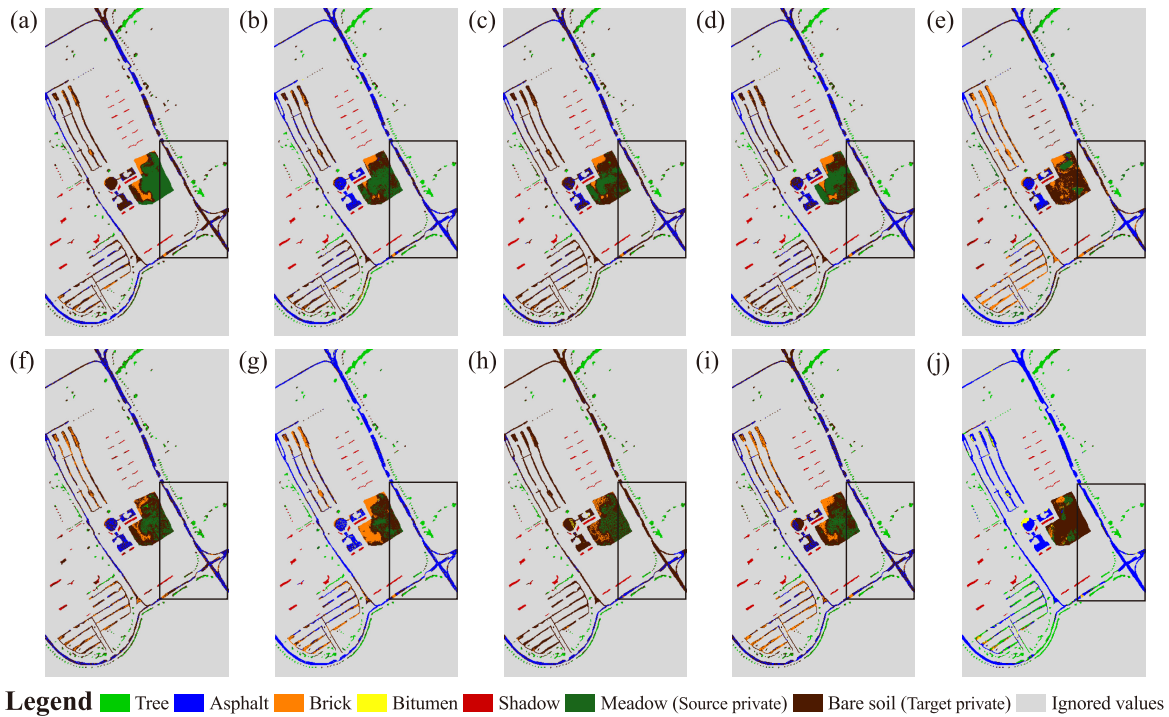

Legend ■ Tree ■ Asphalt ■ Brick ■ Bitumen ■ Shadow ■ Meadow (Source private) ■ Bare soil (Target private) ■ Ignored values

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

proficiency in capturing target-domain-specific information establishes HyUniDA as a robust and effective approach for cross-scene HSI classification tasks.

C. Ablation Studies

For validation and assessment of the HyUniDA's efficacy, we conducted a series of ablation experiments aimed at gaining
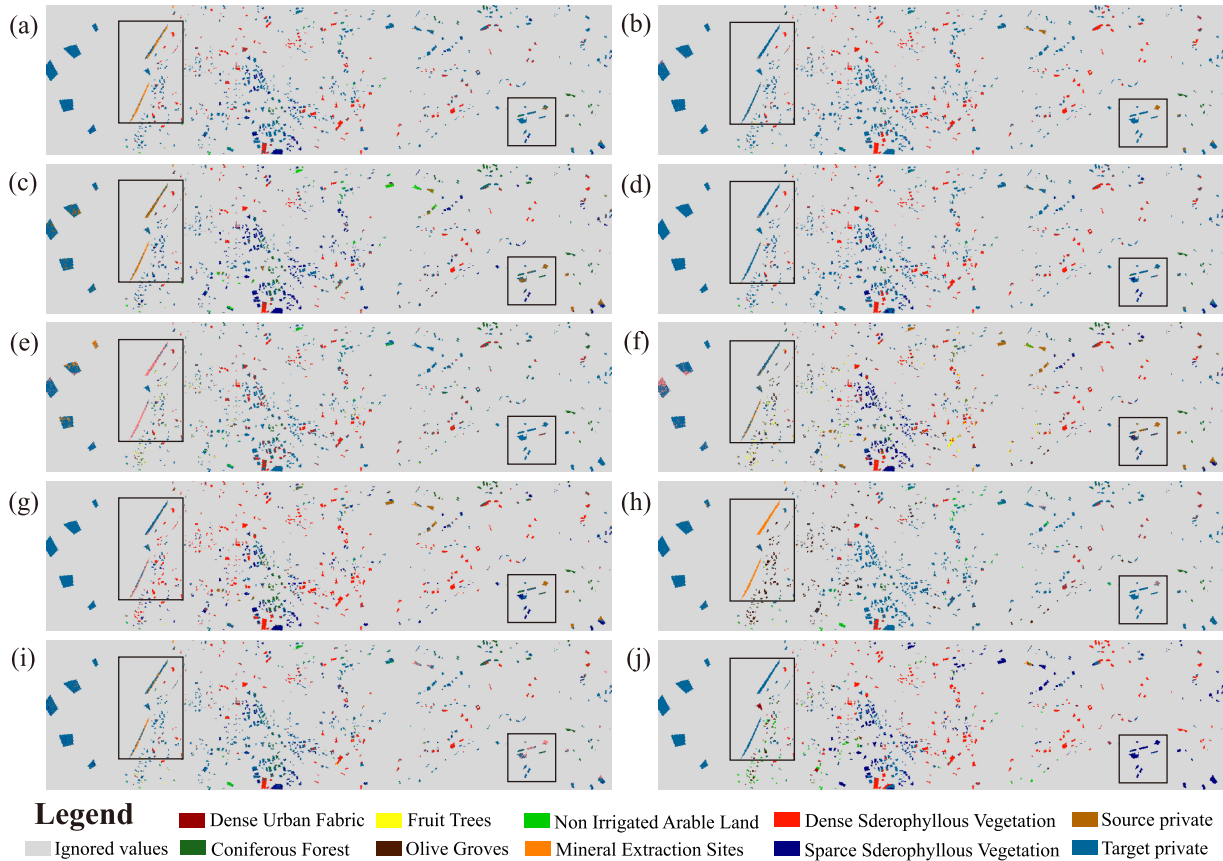
Legend

	Dense Urban Fabric		Fruit Trees		Non Irrigated Arable Land		Dense Sderophyllous Vegetation		Source private
	Ignored values		Coniferous Forest		Olive Groves		Mineral Extraction Sites		Sparce Sderophyllous Vegetation		Target private

Fig. 9. Classification maps from source domain for the target scenario HyRANK Loukia produced by multiple methods, encompassing: (a) Baseline; (b) DANN; (c) CDAN; (d) PADA; (e) OSBP; (f) UAN; (g) DANCE; (h) UniOT; (i) MA; and (j) HyUniDA (Ours). Source private refers to *Sparsely Vegetated Areas* or *Rocks and Sand*, and target private refers to *Water* or *Coastal Water*.

in-depth insights into the contributions of individual components to overall performance. These experiments encompass the impacts of SSP, the effect of different loss functions, sensitivity to hyperparameters, and the robustness in the realistic UniDA setting. Sections IV-C1–IV-C4 will provide a detailed presentation of results from each ablation experiment and their implications for the performance of the proposed model.

1) Effectiveness of the SSP: We delve into the investigation of the SSP mechanism and its influence on the optimization process. Fig. 10 illustrates the dynamic evolution of the number of clusters (K) during the training process across three distinct scenarios. These experiments intentionally omit the use of the proposed stopping criteria outlined in Section II-C. Fig. 10 depicts the convergence process of the cluster numbers, ultimately stabilizing at an optimal value through a few initial attempts. This convergence pattern aligns with the behavior in the DSS evolution. This finding indicates that the exploration of K is crucial in the early stages of training; our method efficiently reaches a stable and optimal value for the number of clusters, which provides empirical support for the efficiency of the stopping criteria offered.

2) Effectiveness of the $\mathcal{L}_{TD}$ and $\mathcal{L}_{CDD}$: The analysis of $\mathcal{L}_{TD}$ and $\mathcal{L}_{CDD}$ contributions involves training the model separately using each term, as detailed in Table VII. When considering simply the cross-entropy loss across source domain ($\mathcal{L}_{SD}$), the model exhibits moderate performance, with accuracy ranging
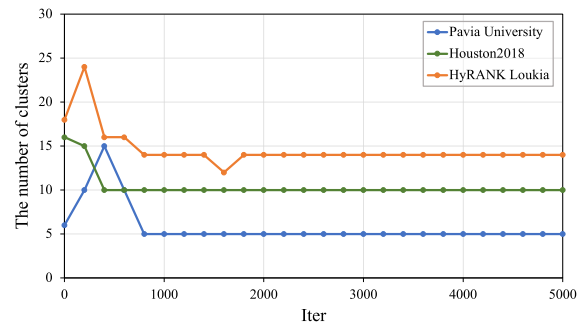

Fig. 10. Evolution of cluster as training progresses.

from 42.68% to 50.88% and H-score from 38.22% to 52.84% across the three tasks. The introduction of $\mathcal{L}_{TD}$ leads to improvements in both accuracy and H-score, indicating the effectiveness of the target-domain loss in aligning the model with the features of the target domain. This enhancement is particularly evident in *TaskC*, where the accuracy jumps from 50.56% to 57.83%, and the H-score increases from 54.25% to 60.77%. The $\mathcal{L}_{CDD}$ further enhances the performance of the HyUniDA by addressing the contrastive domain discrepancies, with improvements in accuracy and H-score observed across almost all tasks. For *TaskB*, the overall accuracy and H-score raised by 4.19% and 3.53%, respectively, after the implementation of $\mathcal{L}_{CDD}$, as opposed to $\mathcal{L}_{TD}$. The best
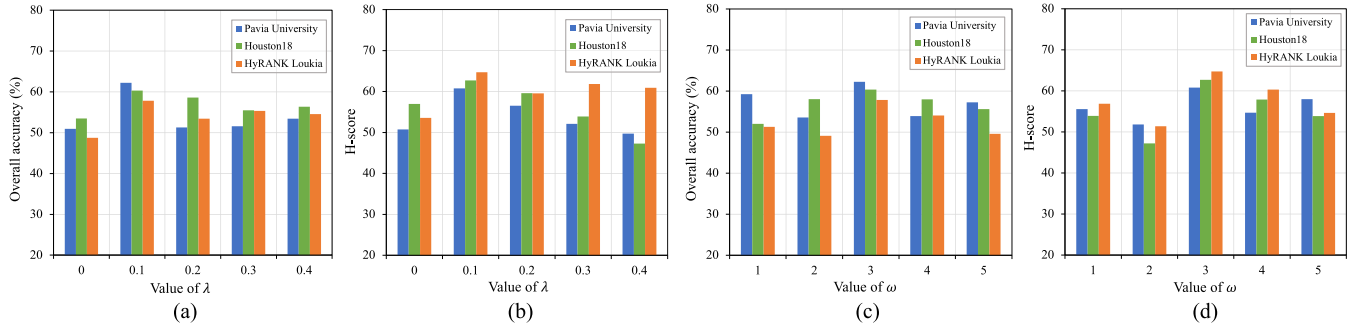
Fig. 11. Sensitivity to hyperparameters. (a) and (b) Impact of λ (ranging from 0 to 0.4) on overall accuracy and H-score, respectively, with fixed $\omega = 3$. (c) and (d) Impact of ω (ranging from 1 to 5) on overall accuracy and H-score, respectively, with fixed $\lambda = 0.1$.

TABLE VII

EVALUATION OF THE EFFECTIVENESS OF THE LOSS COMPONENT, WHERE *TaskA* IS FROM SOURCE SCENE HOUSTON2013 TO TARGET SCENE HOUSTON2018; *TaskB* IS FROM PAVIA CENTER TO PAVIA UNIVERSITY; AND *TaskC* IS FROM HYRANK DIONI TO LOUKIA

$\mathcal{L}_{SD}$	$\mathcal{L}_{TD}$	$\mathcal{L}_{CDD}$	Overall accuracy			H-score		
			TaskA	*TaskB*	*TaskC*	*TaskA*	*TaskB*	*TaskC*
✓			50.88	44.53	42.68	52.84	38.22	48.15
✓	✓		53.48	50.96	48.73	56.96	50.72	53.56
✓		✓	54.52	55.15	50.56	56.43	54.25	55.95
✓	✓	✓	**60.33**	**62.24**	**57.83**	**62.71**	**60.77**	**64.72**

TABLE VIII

ROBUSTNESS IN REALISTIC UNIDA WITH DIVERSE RATIOS OF COMMON CLASSES FOR THE TARGET SCENE HYRANK LOUKIA DATA. DIFFERENT CLASS SPLITS $\mathcal{C}/\overline{\mathcal{C}}_s/\overline{\mathcal{C}}_t$ REVEAL DIVERSE RATIOS OF COMMON CLASSES

	Class spilt					
	7/2/3	7/3/2	8/2/2	9/1/2	9/2/1	10/1/1
Common	45.69	48.43	47.84	48.14	45.38	52.84
Target private	85.18	100.00	100.00	100.00	100.00	100.00
Overall accuracy	57.54	61.60	57.83	57.67	55.42	61.88
H-Score	59.48	65.26	64.72	64.99	62.43	69.15

performance for all tasks is achieved when all three loss terms are turned on, indicating that they are complementary and beneficial for the UniDA scenarios. This comprehensive analysis emphasizes the benefits of incorporating target-domain loss and CDD loss alongside the source-domain loss, affirming their effectiveness in promoting DA performance across diverse HSI classification tasks.

3) Sensitivity to Hyperparameters: We carry out sensitivity analysis on two crucial hyperparameters, namely, λ and ω. Fig. 11 depicts the performance of the overall accuracy and H-score corresponding to the variations of these hyperparameters. Setting ω to 3, we assess the overall accuracy and H-score of parameter λ across a range of 0–0.4, as depicted in Fig. 11(a) and (b), respectively. As can be noticed, the configuration with $\lambda = 0.1$ outperforms alternative settings, demonstrating a significant enhancement. When the value of λ exceeds 0.3, the H-score experiences a sharp decline, except for the HyRANK dataset. For λ fixed at 0.1, the overall accuracy and H-score of ω are assessed across a range of 1–5, as depicted in Fig. 11(c) and (d), respectively. When adjusting the value of ω, the overall accuracy remains relatively stable. However, when $\omega = 2$, it significantly deteriorates the H-score. It should be noted that these two hyperparameters exert varying degrees of influence on each test HSI dataset. Based on our findings, we set up an optimal set of hyperparameters for all our tasks: $\lambda = 0.1$ and $\omega = 3$.

4) Robustness in Realistic UniDA: In a real cross-domain HSI classification scenario, the sample class of the target domain is completely unknown, thus resulting in a multitude of class splits. To assess the robustness of our proposed HyUniDA in realistic UniDA, we carry out experiments with the HyRANK dataset, considering diverse proportions of common classes. Table VIII presents an overview of the findings, where $\mathcal{C}$ represents the number of common classes,

and $\overline{\mathcal{C}}_s$ and $\overline{\mathcal{C}}_t$ are the number of private categories for source and target domains, respectively. It is observed that HyUniDA consistently maintains high overall accuracy and H-score across different splits, showcasing robust generalization capabilities toward shared classes. The overall accuracy ranges from 57.54% to 61.88%, and the H-score varies from 59.48% to 69.15%. Regarding target private classes, HyUniDA exhibits outstanding performance, with the accuracy of more than 85.18%. This means that HyUniDA is very effective for feature learning in the target domain and can accurately distinguish target private samples. Unfortunately, when $\overline{\mathcal{C}}_t \geq 2$, the target domain contains numerous private samples unseen in the source domain. The model struggles to accurately predict each private category, and so only the mean accuracy of target private classes can be presented. In general, the proposed HyUniDA achieves significant overall accuracy and H-score under various class splits, which demonstrated its robustness and excellent performance in handling real UniDA scenarios.

V. DISCUSSION

The HyUniDA we proposed in cross-scene HSI classification addresses a crucial and practical challenge in the remote sensing community, specifically in scenarios where the label set for target domain becomes completely unknown, i.e., private classes may exist in both domains. This section focuses on TE of the practical application scenarios and limitations.

A. TE of Practical Application Scenarios

UniDA poses a pragmatic and intricate transfer challenge. In real-world scenarios, a multitude of situations exists where the source and target samples exhibit inconsistencies. Examples include the challenge of using models trained in urban areas to map land cover in rural landscapes, where there are private samples in both domains (e.g., buildings and

infrastructure in urban areas, and cropland in rural areas). We define the TE as the effectiveness disparity of DA models compared to the Baseline

$$TE = \text{H-score}_{DA} - \text{H-score}_{Baseline}. \qquad (17)$$

TE < 0 denotes that the DA method exhibits a negative TE in the UniDA scenario without any improvement. Notably, when confronted with scenarios where the target-domain label set remains completely undisclosed, traditional DA methods may struggle, leading to unsatisfactory performance. Tables IV–VI indicate that the existing methods have some negative transfer phenomena, such as CDAN (-0.1%) and PADA (-0.44%) in the Houston dataset, and DANN (-0.11%), OSBP (-8.36%), and UniOT (-3.08%) in the HyRANK dataset. For HyRANK datasets, the types of ground objects are complex, and the land cover is diverse, so the spectral differences between Dioni and Loukia may be quite large. Spectral and spatial differences between the source domain and the target domain lead to feature shifts. In the UniDA scenario, the label set of the target domain is unknown, which makes the closed-set DA methods or UniDA methods for general datasets perform poorly when dealing with unlabeled target HSI data. Both the source domain and the target domain have private samples, making the model difficult to generalize. Our newly introduced HyUniDA stands out among other SOTA DA methods, achieving remarkable advancements in handling the issue of cross-scene HSI classification and filling the gap of UniDA in the remote sensing community.

HyUniDA breaks the label set constraints for the first attempt to tackle the UniDA scenario, specifically from HSIs. It consists of the SSP and DSS. SSP effectively identifies cluster pairs with coincident semantic features, allowing for more accurate identification of common classes across different scenes. This innovation addresses the inherent challenge of inconsistent label spaces in cross-scene HSI classification. The introduction of DSS is a unique aspect of HyUniDA that contributes significantly to its performance. DSS not only estimates the number of target clusters but also generates distinct clusters without prior knowledge. This capability enhances the model's adaptability in HSI classification, particularly when dealing with datasets characterized by varying degrees of complexity and diversity. By incorporating CDD, HyUniDA effectively alleviates sample distribution offsets. This aspect is crucial in HSI classification, where differences in sample distributions across different scenes can significantly impact classification accuracy. The regularizer introduced for effective target-domain cluster distinction further enhances the robustness and efficacy of the proposed framework.

B. Limitations and Future Works

While HyUniDA exhibits promising performance, it is essential to address its limitations for a more comprehensive understanding. One constraint lies in its sensitivity to hyperparameters. The effectiveness of HyUniDA may be affected by the chosen hyperparameters, requiring fine-tuning for optimal performance. Moreover, the HyUniDA may exhibit inadequate performance in cases where the feature shift between the source and target domains is significant, or the label distribution is extremely unbalanced. Our model performed poorly for *Fruit Trees* and *Coniferous Forest* in the HyRANK datasets, as did other SOTA methods like DANCE (5.06%) and UniOT (7.59%) for *Fruit Trees*. The MA model designed for remote sensing has low classification accuracy on many categories, such as HyRANK classes 1, 2, and 5. The accuracy of previous SOTA methods is below expectations for the *Brick* of Pavia dataset. UniOT achieved only 2.04%, CDAN 5.13%, and PADA 6.57%.

Future works should focus on refining the method's sensitivity to hyperparameters, exploring methods to automate hyperparameter tuning. Augmenting training data and balancing class representation can mitigate the impact of data imbalance. Strategies like oversampling minority classes or introducing synthetic samples may enhance model robustness. Employing DA techniques tailored for HSI could tackle challenges from domain shifts. Techniques involve domain-invariant feature learning or domain adversarial training. A more in-depth analysis of the spectral features associated with low-accuracy classes may reveal specific challenges. Additionally, strategies to improve recognition for rare private classes in the target domain should be a priority for further research. These efforts will contribute to enhancing the effectiveness and applicability of HyUniDA in diverse and challenging scenarios.

VI. CONCLUSION

In this article, we are the first to discuss cross-scene HSI classification under UniDA scenario, where we transfer knowledge among domains without restrictions of the label space. We propose a UniDA framework based on the SSP and DSS for cross-scene HSI classification, namely, HyUniDA. The SSP identifies cluster pairs with similar semantic features as the common classes, and the DSS calculates the number of target cluster and discovers the private classes from the sample level. Without any preset threshold values, the inherent difference between common and private classes can be automatically detected based on the DSS values of the cluster pairs acquired from the SSP. Furthermore, the CDD is utilized to minimize any distributional bias in the samples, while a prototype regularizer is adopted to promote the separation of target-domain clusters. Our HyUniDA has demonstrated superior performance across three transfer learning tasks for six typical HSI datasets, substantiating its dominance over a wide range of SOTA algorithms.

REFERENCES

[1] Y. Zhang, W. Li, R. Tao, J. Peng, Q. Du, and Z. Cai, "Cross-scene hyperspectral image classification with discriminative cooperative alignment," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 11, pp. 9646–9660, Nov. 2021.

[2] X. Kang, Z. Wang, P. Duan, and X. Wei, "The potential of hyperspectral image classification for oil spill mapping," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–15, 2022, Art. no. 5538415, doi: 10.1109/TGRS.2022.3205966.

[3] R. Näsi et al., "Remote sensing of bark beetle damage in urban forests at individual tree level using a novel hyperspectral camera from UAV and aircraft," *Urban Forestry Urban Greening*, vol. 30, pp. 72–83, Mar. 2018.

[4] J. Liu et al., "Estimating the forage neutral detergent fiber content of alpine grassland in the Tibetan Plateau using hyperspectral data and machine learning algorithms," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–17, 2022, Art. no. 4405017, doi: 10.1109/TGRS.2021.3105482.

[5] 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.

[6] 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.

[7] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2017, pp. 4700–4708.

[8] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jul. 2018, pp. 7132–7141.

[9] W. Huang, Y. Shi, Z. Xiong, Q. Wang, and X. X. Zhu, "Semi-supervised bidirectional alignment for remote sensing cross-domain scene classification," *ISPRS J. Photogramm. Remote Sens.*, vol. 195, pp. 192–203, Jan. 2023.

[10] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1137–1149, Jun. 2017, doi: 10.1109/TPAMI.2016.2577031.

[11] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 779–788.

[12] H. Law and J. Deng, "CornerNet: Detecting objects as paired keypoints," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2018, pp. 734–750.

[13] J. Zheng et al., "Growing status observation for oil palm trees using unmanned aerial vehicle (UAV) images," *ISPRS J. Photogramm. Remote Sens.*, vol. 173, pp. 95–121, Mar. 2021.

[14] R. Zhu and L. Zhuang, "Unsupervised infrared small-object-detection approach of spatial–temporal patch tensor and object selection," *Remote Sens.*, vol. 14, no. 7, p. 1612, Mar. 2022.

[15] S. Minaee, Y. Boykov, F. Porikli, A. Plaza, N. Kehtarnavaz, and D. Terzopoulos, "Image segmentation using deep learning: A survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 7, pp. 3523–3542, Jul. 2021.

[16] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A deep convolutional encoder–decoder architecture for image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 12, pp. 2481–2495, Dec. 2017.

[17] Z. Tian, H. Zhao, M. Shu, Z. Yang, R. Li, and J. Jia, "Prior guided feature enrichment network for few-shot segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 2, pp. 1050–1065, Aug. 2020.

[18] W. Huang, Y. Shi, Z. Xiong, and X. X. Zhu, "AdaptMatch: Adaptive matching for semisupervised binary segmentation of remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, pp. 1–16, 2023, Art. no. 5625416, doi: 10.1109/TGRS.2023.3332490.

[19] N. Wambugu et al., "Hyperspectral image classification on insufficient-sample and feature learning using deep neural networks: A review," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 105, Dec. 2021, Art. no. 102603.

[20] S. Li, W. Song, L. Fang, Y. Chen, P. Ghamisi, and J. A. Benediktsson, "Deep learning for hyperspectral image classification: An overview," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 9, pp. 6690–6709, Sep. 2019.

[21] L. He, J. Li, C. Liu, and S. Li, "Recent advances on spectral–spatial hyperspectral image classification: An overview and new guidelines," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 3, pp. 1579–1597, Mar. 2017.

[22] H. Wu and S. Prasad, "Semi-supervised deep learning using pseudo labels for hyperspectral image classification," *IEEE Trans. Image Process.*, vol. 27, no. 3, pp. 1259–1270, Mar. 2018.

[23] Y. Chen, S. Song, S. Li, L. Yang, and C. Wu, "Domain space transfer extreme learning machine for domain adaptation," *IEEE Trans. Cybern.*, vol. 49, no. 5, pp. 1909–1922, May 2019.

[24] X. Ma, X. Mou, J. Wang, X. Liu, H. Wang, and B. Yin, "Cross-data set hyperspectral image classification based on deep domain adaptation," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 12, pp. 10164–10174, Dec. 2019.

[25] A. Storkey et al., "When training and test sets are different: Characterizing learning transfer," *Dataset Shift Mach. Learn.*, vol. 30, nos. 3–28, p. 6, 2009.

[26] 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.

[27] Y. Qin, L. Bruzzone, and B. Li, "Tensor alignment based domain adaptation for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 11, pp. 9290–9307, Nov. 2019.

[28] S. Nirmal, V. Sowmya, and K. Soman, "Open set domain adaptation for hyperspectral image classification using generative adversarial network," in *Proc. Inventive Commun. Comput. Technol. (ICICCT)*. Cham, Switzerland: Springer, 2020, pp. 819–827.

[29] J. Zheng, W. Wu, S. Yuan, H. Fu, W. Li, and L. Yu, "Multisource-domain generalization-based oil palm tree detection using very-high-resolution (VHR) satellite images," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022.

[30] Q. Zhu et al., "Land-use/land-cover change detection based on a Siamese global learning framework for high spatial resolution remote sensing imagery," *ISPRS J. Photogramm. Remote Sens.*, vol. 184, pp. 63–78, Feb. 2022.

[31] W. M. Kouw and M. Loog, "A review of domain adaptation without target labels," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 3, pp. 766–785, Mar. 2021.

[32] R. Zhu, L. Yan, N. Mo, and Y. Liu, "Semi-supervised center-based discriminative adversarial learning for cross-domain scene-level land-cover classification of aerial images," *ISPRS J. Photogramm. Remote Sens.*, vol. 155, pp. 72–89, Sep. 2019.

[33] X. Liu et al., "Adversarial unsupervised domain adaptation with conditional and label shift: Infer, align and iterate," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2021, pp. 10367–10376.

[34] 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, pp. 1–13, 2022, Art. no. 5609213, doi: 10.1109/TGRS.2021.3105302.

[35] H. Yan, Y. Ding, P. Li, Q. Wang, Y. Xu, and W. Zuo, "Mind the class weight bias: Weighted maximum mean discrepancy for unsupervised domain adaptation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2017, pp. 2272–2281.

[36] Z. Pei, Z. Cao, M. Long, and J. Wang, "Multi-adversarial domain adaptation," in *Proc. AAAI*, Apr. 2018, vol. 32, no. 1, pp. 3934–3941, doi: 10.1609/aaai.v32i1.11767.

[37] M. Long, Z. Cao, J. Wang, and M. I. Jordan, "Conditional adversarial domain adaptation," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 31, 2018.

[38] 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, doi: 10.1109/TNNLS.2021.3109872.

[39] Y. Zhang, W. Li, W. Sun, R. Tao, and Q. Du, "Single-source domain expansion network for cross-scene hyperspectral image classification," *IEEE Trans. Image Process.*, vol. 32, pp. 1498–1512, 2023.

[40] C. Deng, X. Liu, C. Li, and D. Tao, "Active multi-kernel domain adaptation for hyperspectral image classification," *Pattern Recognit.*, vol. 77, pp. 306–315, May 2018.

[41] H. Wang, X. Wang, C. L. P. Chen, and Y. Cheng, "Hyperspectral image classification based on domain adaptation broad learning," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 3006–3018, 2020.

[42] 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.

[43] J. Yue, L. Fang, and M. He, "Spectral–spatial latent reconstruction for open-set hyperspectral image classification," *IEEE Trans. Image Process.*, vol. 31, pp. 5227–5241, 2022.

[44] Z. Xie, P. Duan, W. Liu, X. Kang, X. Wei, and S. Li, "Feature consistency-based prototype network for open-set hyperspectral image classification," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Jan. 6, 2023, doi: 10.1109/TNNLS.2022.3232225.

[45] 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.

[46] Z. Cao, L. Ma, M. Long, and J. Wang, "Partial adversarial domain adaptation," in *Proc. Eur. Conf. Comput. Vis.*, 2018, pp. 135–150.

[47] 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, pp. 1–17, 2023, Art. no. 5601317, doi: 10.1109/TGRS.2022.3229039.

[48] 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.

[49] G. Li, G. Kang, Y. Zhu, Y. Wei, and Y. Yang, "Domain consensus clustering for universal domain adaptation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 9757–9766.

[50] J. Guo et al., "C³DA: A universal domain adaptation method for scene classification from remote sensing imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 21, pp. 1–5, 2024.

[51] K. Saito, D. Kim, S. Sclaroff, and K. Saenko, "Universal domain adaptation through self supervision," in *Proc. NIPS*, vol. 33, 2020, pp. 16282–16292.

[52] W. Chang, Y. Shi, H. Tuan, and J. Wang, "Unified optimal transport framework for universal domain adaptation," in *Proc. Adv. Neural Inf. Process. Syst.*, 2022, pp. 29512–29524.

[53] 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, pp. 1–15, 2023, Art. no. 4700515, doi: 10.1109/TGRS.2023.3235988.

[54] Y. Zhang, W. Li, M. Zhang, S. Wang, R. Tao, and Q. Du, "Graph information aggregation cross-domain few-shot learning for hyperspectral image classification," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 2, pp. 1912–1925, Feb. 2024, doi: 10.1109/TNNLS.2022.3185795.

[55] Z. Fang et al., "Confident learning-based domain adaptation for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5527116.

[56] A. E. Ezugwu et al., "A comprehensive survey of clustering algorithms: State-of-the-art machine learning applications, taxonomy, challenges, and future research prospects," *Eng. Appl. Artif. Intell.*, vol. 110, Apr. 2022, Art. no. 104743.

[57] A. Gretton, K. Borgwardt, M. Rasch, B. Schölkopf, and A. Smola, "A kernel method for the two-sample-problem," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 19, 2006.

[58] G. Kang, L. Jiang, Y. Yang, and A. G. Hauptmann, "Contrastive adaptation network for unsupervised domain adaptation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2019, pp. 4893–4902.

[59] G. Kang, L. Jiang, Y. Wei, Y. Yang, and A. Hauptmann, "Contrastive adaptation network for single- and multi-source domain adaptation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 4, pp. 1793–1804, Apr. 2022.

[60] 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.

[61] B. Le Saux, N. Yokoya, R. Hansch, and S. Prasad, "2018 IEEE GRSS data fusion contest: Multimodal land use classification [technical committees]," *IEEE Geosci. Remote Sens. Mag.*, vol. 6, no. 1, pp. 52–54, Mar. 2018.

[62] K. Karantzalos, C. Karakizi, Z. Kandylakis, and G. Antoniou, "HyRANK hyperspectral satellite dataset I (version v001)," Int. Soc. Photogramm. Remote Sens., Tech. Rep., 2018, doi: 10.5281/zenodo.1222202.

[63] B. Fu, Z. Cao, M. Long, and J. Wang, "Learning to detect open classes for universal domain adaptation," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*. Cham, Switzerland: Springer, 2020, pp. 567–583.

[64] Y. Ganin et al., "Domain-adversarial training of neural networks," *J. Mach. Learn. Res.*, vol. 17, no. 1, pp. 2030–2096, Apr. 2016.

[65] K. Saito, S. Yamamoto, Y. Ushiku, and T. Harada, "open set domain adaptation by backpropagation," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2018, pp. 153–168.