Unsupervised Domain Adaptation With Dense-Based Compaction for Hyperspectral Imagery

Chunyan Yu , *Member, IEEE*, Caiyu Liu, *Member, IEEE*, Haoyang Yu , *Member, IEEE*,
Meiping Song, *Member, IEEE*, and Chein-I Chang , *Life Fellow, IEEE*

Abstract—**Enormously hard work of label obtaining leads to the lack of enough annotated samples in the hyperspectral imagery (HSI). The mentioned reality inferred the unsupervised classification performance barely satisfactorily. Unsupervised domain adaptation is exploited for knowledge delivery from a labeled source domain to boost the performance on an unlabeled target domain. In this article, we propose an unsupervised domain adaptation architecture with dense-based compaction (UDAD) for HSI classification (HSIC). The processes of spectral–spatial feature compaction, unsupervised domain adaptation, and classifier training are incorporated with an integrated framework to complete the HSI cross-scene classification. The core of the proposed framework is to utilize adversarial domain learning to reduce the domain discrepancy. To this end, the classifier trained in the source domain would accomplish well in the target domain for the unsupervised HSIC. Besides, to extract the discriminative spectral–spatial feature for the HSI domains, a dense-based compaction network is applied to complete the semisymmetric mapping. Our experiments illustrate that the UDAD model yields more effective classification performance than other state-of-the-art unsupervised HSIC methods.**

Index Terms—**Adversarial training, domain adaptation, HSI classification (HSIC), transfer learning.**

I. INTRODUCTION

HYPERSPECTRAL remote sensing technology benefits from the development of spectral imaging and supplies sufficient information for target recognition through hundreds of continuous and subdivided spectral bands [1], [2]. As a typical and fundamental application of the hyperspectral imagery (HSI) analysis, HSI classification (HSIC) plays a significant role in the fields of land cover detection and agriculture

Chunyan Yu, Caiyu Liu, Haoyang Yu, and Meiping Song are with the Center for Hyperspectral Imaging in Remote Sensing (CHIRS), Dalian Maritime University, Dalian 116026, China (e-mail: yuchunyan1997@126.com; liu_caiyu@163.com; yuhy@dlmu.edu.cn; smping@163.com).

Chein-I Chang is with the Center for Hyperspectral Imaging in Remote Sensing (CHIRS), Dalian Maritime University, Dalian 116026, China, and also with the Remote Sensing Signal and Image Processing Laboratory, Department of Computer Science and Electrical Engineering, University of Maryland, Baltimore, MD 21250 USA (e-mail: cchang@umbc.edu).

monitoring [3]–[9]. Nowadays, the approaches utilizing spatial features assisted with spectral information to improve the classification performance have attracted considerable attention in the fields of HSIC [10]–[14]. In [10], an iterative target-constrained interference-minimization classifier performed spectral–spatial HSIC by adopting an iterative Gaussian filtered feedback mode to combine the spatial context information to spectral feature. In [13], the HSIC model based on sparse theory was established, which achieved excellent performance with the sparse representation of the HSI sample. With the progress of the various artificial neural network structure, spectral–spatial classification models based on deep learning have made breakthroughs in the HSIC fields [15]–[26]. Yu *et al.* [15] proposed a 2-D-CNN architecture with a deconvolution layer and embedded hashing semantic features to improve the HSIC accuracy. In [16], an in-depth feature learning model was presented based on a multidecision labeling policy, which takes two decision measures to prelabel local and global samples. In recent years, graph convolutional network (GCN) aims for irregular data representation, and analysis has been applied in the processing of hyperspectral data. In [17], a multiscale dynamic graph convolution framework was presented to reduce the impact of a low-quality predefined graph, which exploited the refined feature representation for irregular HSI cube. Hong *et al.* [18], proposed supervised miniGCNs in a minibatch fashion to overcome the barrier of substantial computational cost, especially in large-scale remote sensing occasions. Meanwhile, the generative adversarial network (GAN) has drawn much attention to generating samples for solving complex labeling problems in hyperspectral datasets. In [19], a GAN-based HSIC system was employed, the adversarial training for sample generation that consisted of two versions: 1-D spectral and 3-D spectral–spatial classifiers. Wang *et al.* [20], developed a semisupervised variational GAN for synthetic sample generation, which adopted an encoder–decoder framework to complete the feature mapping. It is worth noting that most current models primarily focus on the HSIC in a supervised style with label annotation.

With a wealth of observation data produced by the hyperspectral satellites, a large number of unlabeled samples cannot be directly used for feature recognition and classification. Therefore, utilizing a vast amount of samples without label annotation, called unsupervised HSIC, meets the current target classification requirement. Compared with the time-consuming and challenging task of labeling samples, cross-scene classification learning

has attracted much attention in HSIC [27]–[29]. Recently, transfer learning provides a popular option to solve the cross-scene classification of HSI, which aims to transfer the knowledge implied in the source domain to the target domain by feature adaption. In this style, the classifier trained with the label sample in the source domain can produce a compelling performance in the target scene. The domain feature is typically aligned on mapping subspace in the traditional domain transferring pattern by embedding kernel subspace based on the shallow structure [30]–[34]. This kind of realization mainly used the projection principle to map the features of the two domains into the same subspace and obtain consistent features through subspace alignment. In [29], an unsupervised representation method was accomplished for HSI scene classification by utilizing a weighted deconvolution network to learn the feature maps. In [30], an HFAA model based on feature adaptation and augmentation for cross-scene HSIC was proposed, which gained a common subspace with projection matrices to extract the transferable knowledge. Qian *et al.* [31] presented a dictionary learning-based domain adaptation for cross-scene HSIC, which was implemented by projecting the spectral features of the source and target domain into a shared embedding space based on multitasking dictionary learning.

Due to the limitations of the modeling and representation capabilities of the shallow structure, the mentioned scheme is difficult to obtain satisfactory classification performance. In recent years, experts and scholars have introduced deep learning into hyperspectral transfer learning, and proposed the deep transfer learning framework to achieve cross-domain feature mapping. This type of approach utilizes deep learning networks to implement feature conversion and distribution alignment, which constructs measurement criteria in the spectral space to make the distribution of the maps of the HSI domains as close as possible. The deep transfer learning model [35]–[37] usually consists of: a) a pertrained network to extract the spectral–spatial feature of the source domain and train the labeled sample to generate a classifier and b) a fine-tuning phase adopted in the target domain to implement the feature adaption. In [35], a two-branch deep CNN architecture was established independently to extract the spectral and spatial features. The lower layers of the proposed network were pretrained in the source domain, and top layers were trained for the target scene classification after fine-tuning. Jiang *et al.* [36] proposed a 3-D-SRNet method to implement HSI cross-scene transfer learning by collaborating the 3-D separable ResNet to extract the spectral–spatial feature instead of the 3-D convolution operation. In which the pretrained model is transferred to the target HSI dataset by fine-tuning. In [37], a cross-scene deep transfer learning method was built on sharing the effective contents with a joint probability distribution adaptation approach. The mentioned deep transfer learning classification methods have brought significant progress to the cross-scene HSIC. Nevertheless, the classification accuracy propagates with the insufficient feature extraction during the hyperspectral source domain pretraining process, which results in negative migration. Besides, the complicated discrepancy between the source domain and target domain caused by the

hyperspectral imager and different sun-view poses challenging issues of domain adaptation with limited training samples. Recently, the GAN framework [38] integrated with a generator and a discriminator has attracted extensive attention in various machine learning tasks. Notably, the adversarial discriminative domain adaptation (ADDA) model [39] based on GAN has supplied a new approach for transfer learning, which is implemented with adversarial training to solve the domain shift for the feature adaptation. Nowadays, unsupervised domain adaptation [40], [41] attempt to handle complicated scenarios without annotated information attracts more attention in the application field. Generally, in an unsupervised domain adaptation, the source and target domains are drawn from different distributions but contain the same categories. The unsupervised pattern is implemented without leveraging the label information of the target scenes.

This article proposes an unsupervised domain adaptation architecture with a dense-based compaction (UDAD) model for HSI transfer learning classification. It needs to be noted that the application assumption of our method is that the source HSI has labeled samples, whereas the target HSI lacks labeled training data for a classification task. Practically, we accomplish the discrepancy reduction with the designed adversarial domain learning framework in the proposed method. In this way, the classifier pretrained with the labeled samples of the source domain is acceptable for the unlabeled target domain. Therefore, we call the unsupervised classification for the target scene. The UDAD framework is mainly composed of three phases, including labeled domain training, unsupervised domain adaptation, and unsupervised target classification. A spectral–spatial dense network with a compact style of feature organization is employed to construct feature representation of the source domain in the learning phase. Adversarial domain adaptation based on the semisymmetric mechanism is built to complete the feature alignment between the source and the target domain. The discriminator in the UDAD model is implemented with the fully connection layer, which is responsible for the recognition of the mapping source. Afterward, the trained classifier is tested in the target domain to complete the HSI target scene classification with an unsupervised mode. In summary, the main contributions of this article are listed as follows.

1) The fine-tuning phase for the traditional adaptation is ineffective due to the lack of labeled samples in our assumption. A new unsupervised adaptation model for HSIC is presented in this article. To the best of our knowledge, the unsupervised domain adaptation implemented with dense feature compaction is the first attempted to develop for feature alignment of HSIC with the cooperation of the compact style and an adversarial discriminator.

2) The backbone of the training network is constructed by a dense-based CNN in a concatenation style. The advantage of this pattern performs discriminative feature representation in the labeled learning phase, which is beneficial for knowledge transferring during the unsupervised domain adaptation.

3) We compared extensive experiments under the different domain discrepancies, and provided a comprehensive
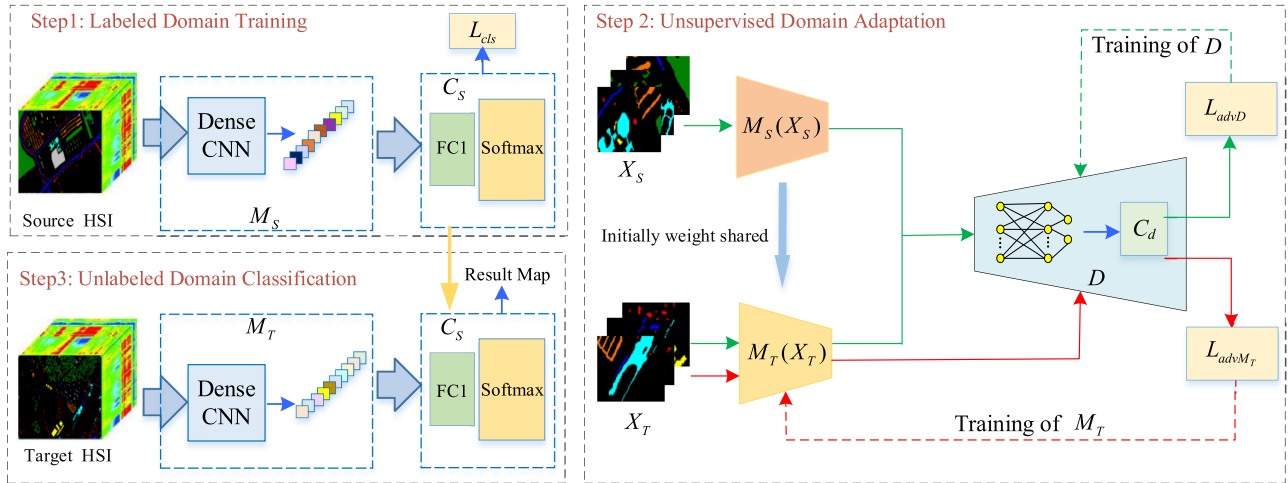
Fig. 1.   Overview of hyperspectral cross-scene classification with the UDAD network. The proposed scheme is conducted in three steps. First, the discriminative spectral–spatial features of the HSI source domain are exploited in the trained phase. Next, the domain distribution alignment is completed with adversarial learning based on the mapping of different scenarios. Finally, the unsupervised HSIC of the target scene is accomplished with the trained classifier.

analysis of the classification results of the UDAD framework, which demonstrated the proposed method could effectively improve the domain generalization ability.

The rest of this article is organized as follows. In Section II, the details of the proposed UDAD architecture are described. Section III provides the experimental results and analysis. Finally, Section IV concludes this article.

II. PROPOSED APPROACH

Fig. 1 illustrates the flowchart of the proposed UDAD model composed of three modules, including source domain pretraining, adversarial domain adaptation, and the target domain classification. First, the source domain training phase is adopted to extract the most discriminative spectral–spatial feature for training the source HSI classifier. Second, an unsupervised domain adaptation focuses on the spectral–spatial feature adaption, which is implemented by a designed generative adversarial framework with a discriminator that attempts to distinguish the feature difference. Finally, the unsupervised classification of the target HSI is predicted by the pretrained classifier of the source domain. The components of the framework are further described in detail.

A. Labeled Domain Training

Practically, assume that the sample of an HSI in the labeled source domain is denoted as X with the dimension of $m \times n \times u$, and the distribution with sample X labeled Y is defined as P_S. Similarly, the distribution of the unlabeled target domain is recorded as P_T. The core task is to acquire a rich representation of the source domain of HSI and improve the discriminate knowledge for transfer learning.

In the UDAD model, we construct the dense convolutional framework denoted as M_S for the spectral–spatial feature extraction of the source HSI. The detailed structure of the M_S is shown in Fig. 2, which is a multilayer network that operates on a dense-based network. Specifically, we employ two dense blocks to assemble discriminative spectral–spatial features. To be specific, each dense block with different kernel sizes consists of two convolution layers and one activation layer.

The fundamental convolution layer is designed to exploit the spectral–spatial feature via sharing weights and biases of neurons in adjacent layers. The produced map f_n performed convolution is formulated by

$$f_n = f_i^n \Theta w + b \tag{1}$$

where ω and b are the weights and biases of the network, Θ refers to the 2-D convolution operator, and f_i^n represents the compacted input of the n*th* layer. Subsequently, the feature mapping is activated by the ReLu function.

The dense block is primarily utilized to concatenate the output of all previous layers in a feed-forward style, aiming to reduce the possibility of gradient disappearance. In this style, the compact network can benefit from more in-depth network training. The dense operation is defined as follows:

$$f_i^n = \sum_{p=1}^{n-1} f_p. \tag{2}$$

Except for the extraction module M_S, the classifier C_S is essential for the labeled source training phase. We design the classifier with one fully connected layer and one activation layer. Based on the training sample of the source domain samples $\{(X_s, Y_s)\} = \{(x_1, y_1), \ldots, (x_n, y_n)\}$, the training of C_S is completed by the categorical cross-entropy loss.

$$L_{cls}(X_S, Y_S) = E_{(x_s, y_s) \sim (X_S, Y_S)} - \sum_{k=1}^{K} I_{[k=y_s]}$$
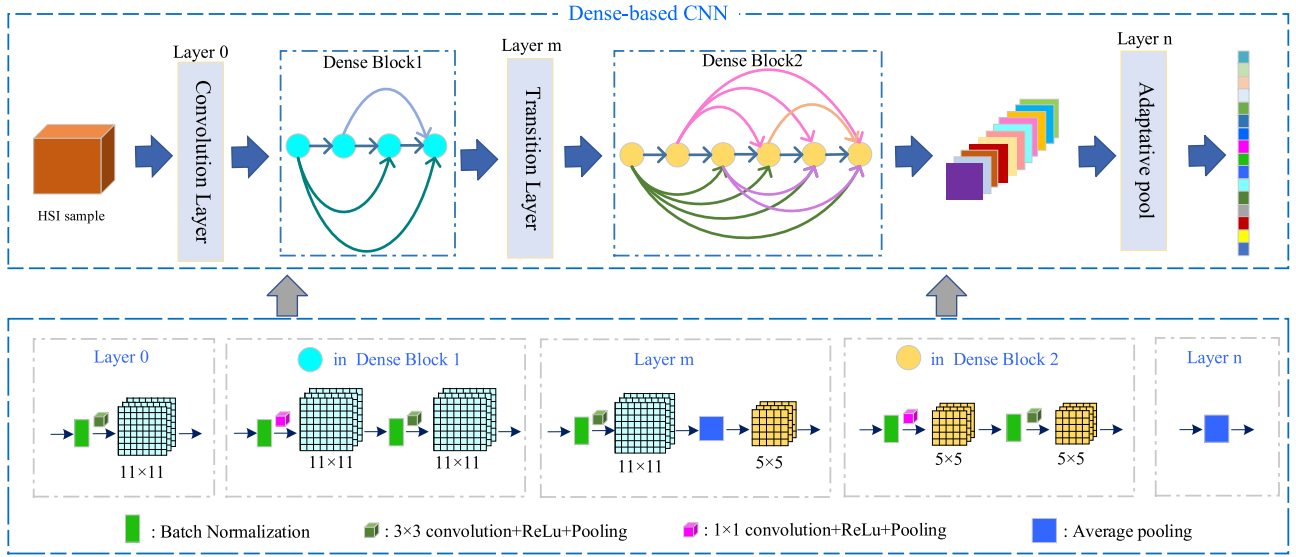$$\times \log C_s(M_S(X_S)) \tag{3}$$

Fig. 2.　Structure of the dense CNN of the proposed UDAD network. The node ∘ denotes a dense module. Three embedding layers are utilized in the compact model, including the convolution layer (Layer 0) for initial spatial feature extraction, the transition layer (Layer m) for feature compression, and the adaptive pooling layer (Layer n) for feature nonlinearization.

where M_S represents the extracted feature, K represents the number of the class, E denotes the expectation, and I represents the identity matrix.

B. Unsupervised Domain Adaptation

Unsupervised domain adaptation for the cross-domain HSIC aims to decrease the domain shift and improve the alignment capability between a labeled source HSI and an unlabeled target HSI. In this article, we proposed an adversarial domain adaptation framework for HSI based on the recently designed ADDA framework to generalize the feature distribution of the source domain M_S and target domain M_T. Fig. 3 shows detail of the adversarial discrimination of the proposed model. In which the M_S and M_T that built by the dense network obtain the spectral–spatial mappings of the two scenarios. Adversarial learning for HSI feature alignment is implemented by alternating minimization between two functions based on GAN loss. In the following part, we give the adversarial training procedure of the proposed method.

The domain discriminator, denoted by D, is used to distinguish which domain the spectral–spatial feature mapping belongs to. As shown in the top branch of Fig. 3, both M_S and M_T are fixed during the D training phase. The convergence capability is improved by first initializing the weights of M_T with the weights of M_S, i.e., $\theta_{M_T} = \theta_{M_S}$. The training sample set is denoted by $X = [X_S^*, X_T]$, where $X^*_S = \{(x_{S_1}), \ldots, (x_{S_n})\}$ is the unlabeled source domain samples, and $X_T = \{(x_{T_1}), \ldots, (x_{T_n})\}$ represent the unlabeled target domain samples. The task of the D is to declare which domain the feature mapping $M_S(X_S^*)$ and $M_T(X_T)$ comes from, which is recorded as C_d. We assign the label of 1 and 0 to the source and target domains, respectively. The discriminator D is optimized
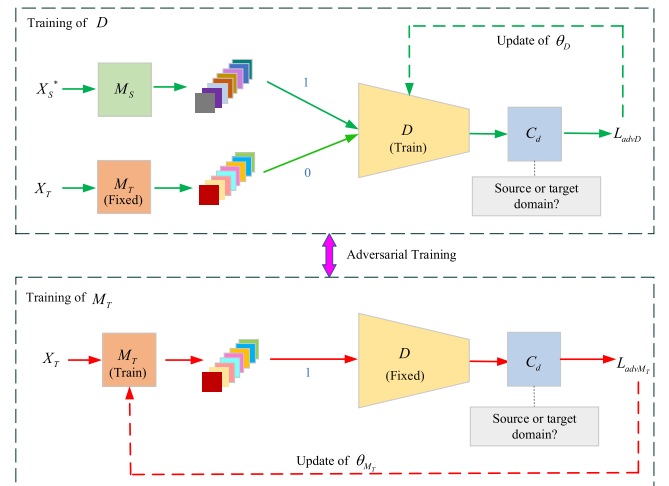

Fig. 3.　Detail of the adversarial training of the unsupervised adaptation part. On the whole, two subsections proceed alternatively to achieve the adaptation. The top figure shows the training implementation of the discriminator specified with the flow of the green line, and the bottom figure demonstrates the training process of the feature extractor of the target domain indicated with the flow of the red line.

according to a standard supervised loss by the following formula:

$$L_{advD}(X_S, X_T, M_S, M_T) = -E_{x_S \sim X_S}[\log D(M_S(X_S))]$$
$$- E_{x_T \sim X_T}[\log(1 - D(M_T(X_T)))]. \quad (4)$$

Adversarially, the M_T need to construct an output that has the same contribution to the output of the M_S. As shown in the bottom branch of Fig. 3, the purpose of M_T optimization is to unify the distribution of the two domains, and we use the defined objective to train M_T alternatively until aligning the target spectral–spatial distribution to the source HSI. The label

of the sample X_T of the target domain is assigned as 1 in the implementation of this phase. The target mapping is optimized with D, fixed according to a constrained adversarial objective for M_T defined as follows:

$$L_{advM_T}(M_T, X_T, D) = -E_{x_T \sim X_T}[\log D(M_T(X_T))]. \quad (5)$$

C. Implementation Details of the UDAD Model

Due to the characteristics of feature reusing, the dense network performs well even with a small number of output layers. However, the compact connection causes an excessive burden with the depth addition. To solve this problem, in this article, we separated the whole dense network into two blocks to decrease the number of the output channel. Moreover, to increase the fitting ability, we utilize a transition layer as the intermediate node to implement feature integration composed of the bottleneck and max-pooling layers. Specifically, the max-pooling layer is employed to accomplish the feature compression. A 1×1 convolution layer is utilized as a bottleneck layer for feature dimensionality reduction in the dense block. In general, the pooling layer and the bottleneck layer are used simultaneously to decrease the parameters of the UDAD. Two fully connected layers with 50 neurons build the discriminator network. Besides solving the collapse problem, the batch normalization policy is applied in the independent module, and the optimizer adopted is ADAM.

The HSI target domain is unlabeled in the adversarial domain adaption, which caused an optimization problem best to minimize the distance between the source and target mappings. Recent methods have favored sharing weights partially to learn parameters for each domain individually. In this work, a semisymmetric mechanism is applied to achieve effective adaptation, in which not all parameters are participated in learning an asymmetric mapping of the feature space. In detail, we initialize the parameters of M_T, the target domain with the trained settings M_S, and the unsupervised adaption learning is implemented by training M_T and D in an adversarial style by minimizing the GAN-based loss.

D. Unsupervised Classification on Target Domain

After the spectral–spatial feature adaption, the different scenarios are provided with a similar feature distribution. Therefore, the classifier for the source domain performed on the target domain exhibits promising performance for HSI cross-scene classification. In addition, the samples in the target domain are unlabeled, which is an unsupervised training mode. In summary, the steps of the unsupervised cross-scene HSIC with the UDAD architecture are outlined as follows.

III. EXPERIMENT AND RESULT ANALYSIS

A. Data Description

In this section, we adopted three HSI datasets to evaluate the performance of the proposed method. The first one is called the Pavia University data, and it is a part of the hyperspectral image

---

Algorithm 1: Training of the Labeled Domain.

Input: $\{(X_s, Y_s)\}$, iteration number N_1.
Output: Weights and bias of M_S and C_S.
Initialize: $\theta_{M_S}, \theta_{C_S}$
For $i = 1$ to N_1 do
  Optimize M_S, C_S according to the (3) with the Gradient descent method.
  $\theta_{MS}, \theta_C = \leftarrow \nabla_{\theta MS, \theta C}(L_{cls})$
End

Algorithm 2: Unsupervised Domain Adaptation.

Input: $X = [X_s^*, X_t]$, iteration number N_2.
Output: Weights and bias of M_T and D.
Initialize: $\theta_D, \theta_{M_S}, \theta_{M_T}$
For $i = 1$ to N_2 do
  Optimize M_T, D in turn according to the (4) and (5) with the Gradient descent method.
  $\theta_D = \leftarrow -\nabla_{\theta_D}(L_{advD})$
  $\theta_{M_T} = \leftarrow -\nabla_{\theta_{M_t}}(L_{advM_T})$
End

Algorithm 3: Unsupervised HSIC of the Target Domain.

Input: $X_T = \{(x_{T_1}), \ldots, (x_{T_N})\}$, Sample number N.
Output: Prediction label $\{Y_{T_1} \cdots Y_{T_N}\}$
For $i = 1$ to N do
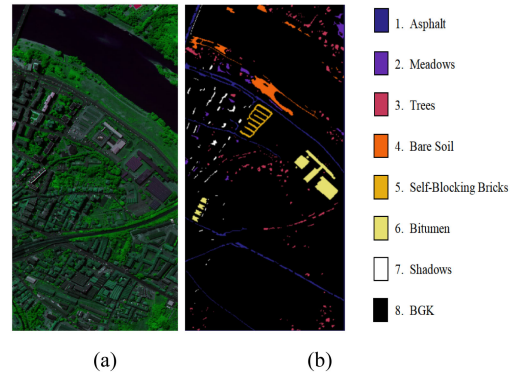 $Y_{T_i} = C_S(M_T(x_{T_i}))$
End

Fig. 4. Pavia Center scene. (a) False-color image with band (13, 78, 41). (b) Colorful ground-truth image.

of Pavia city in Italy made. This image contains 103 bands with a wavelength range of 0.43–0.86 μm after removing the noise and water bands, and the spatial resolution is 1.3 m. The data consists of nine types of classes with a resolution size 610×340. In the following experiment, we analyze this scene together with the Pavia Center to accomplish the cross-scene knowledge transferring. The last band of the original cube is abandoned, to keep the same band number with the other Pavia scene. A false-color image is shown in Fig. 4(a).
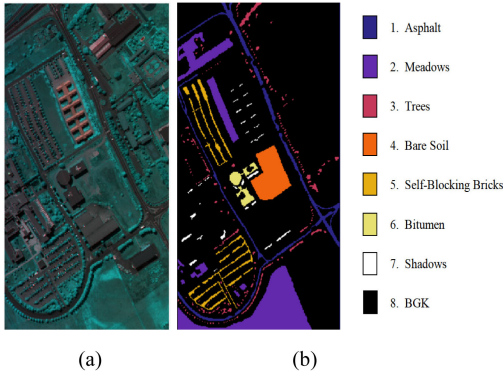
Fig. 5.   Pavia University scene. (a) False-color image with band (34, 97, 85). (b) Colorful ground-truth image.

TABLE I
HYPERPARAMETERS SETTING OF THE HS-ADDA

Layer	Input	Kernel sets
Convolution 1 with the same padding	$11\times11\times$channel	3×3, 324
ReLu+Max Pooling	$11\times11\times324$	2×2, Stride=2
Convolution 2 with the same padding	$5\times5\times324$	3×3, 432
Global Ave Pooling	$5\times5\times432$	—
Fully connection 1	$1\times1\times432$	100
Fully connection 2	$1\times1\times100$	50
Softmax	$1\times1\times50$	—

TABLE II
HYPERPARAMETERS SETTING OF THE DENSE CNN FOR FEATURE EXTRACTION

Module	Input	Pooling
Layer 0	$11\times11\times$channel	MaxPool
Dense block 1	$11\times11\times96$	—
	$11\times11\times144$	—
	$11\times11\times192$	—
	$11\times11\times240$	—
Layer m	$11\times11\times288$	AvgPool
Dense block 2	$5\times5\times144$	—
	$5\times5\times192$	—
	$5\times5\times240$	—
	$5\times5\times288$	—
	$5\times5\times336$	—
	$5\times5\times384$	—
Layer n	$5\times5\times432$	AvgPool

The second dataset is the Pavia Center, which was captured by ROSIS with the same spectral coverage and spatial resolution as the Pavia University. The image is 1096×492 with 102 spectral bands. Although the two datasets of Pavia have nine classes, the seven same categories are analyzed in our experiments. A false-color image is shown in Fig. 5(a), the ground truth image of the Pavia University and the Pavia Center with seven classes are shown in Figs. 4(b) and 5(b), respectively.

The last dataset is the Botswana data, which was acquired by the NASA EO-1 Satellite over the Okavango Delta in the period 2001–2004. This scene consists of 256×1476 pixels with 145 bands, after removing noisy bands. In our experiments, we have divided the image into the top and bottom parts as the source and target domain, respectively. The false-color composite image and the ground truth image with six picked classes are shown in Fig. 6(a) and (b).

B. Experimental Settings

To examine how the proposed architecture benefits the HSIC of the target domain without training labels, we exploited a series of other unsupervised classification models for comparison with the UDAD. In our article, KNN is implemented with principal component analysis and k-nearest neighbor cluster algorithm. The OSDA approach [42] was implemented by open set domain adaptation and GAN, which created a cooperative model to classify data from different classes. For a fair comparison, we implemented the OSDA method with the same hyperparameters as shown in [42], except for the batch size of 64 and iteration of 300. The DANN method [43] achieved domain adaptation by aligning the distributions of features through standard backpropagation training. In our experiment, we set the kernel size of the feature extraction network to 3×3 instead of 5×5, and the classifier and the discriminator have the same setting as the reference.

The unsupervised classification method with the classifier trained from the source domain is recorded as the UDHS method, which is implemented without any domain adaption. Whereas, we accomplish the discrepancy reduction with the designed adversarial domain learning framework in UDAD. The other compared methods are the deformation of the proposed model,

including the HS-D1 model for the network UDAD without the second dense block, and the HS-D2 method for the proposed system without the first dense block. The UDAD without the compact scheme is recorded as HS-ADDA [40]. The settings of the CNN-based feature extractor are given in Table I. All experimental results are performed on a computer with a 64-bit Windows10 system, which is configured with Intel (R) Pentium (R) G4560 SR32Y@ 3.50GHz, 16G operating memory, PyTorch-Cuda 10-cudnn7 1.4.0, and python 3.6 simulation platform. We adopt the overall accuracy (OA), average accuracy (AA), and statistic kappa coefficient (Kappa) as objective criteria to evaluate the performance of each method.

The execution of all the compared methods is conducted by five-times. All results reported in our experiment are recorded in the form of mean $\pm$ standard deviation. The learning rate is set to 0.0001, the patch size is 11×11, N_1 for the source classifier is set to 200, and N_2 is equal to 1000 for the adversarial training. Further, according to the semisymmetric mechanism, the same configuration of M_S and M_T is designated in Table II. Table III gives the number of training samples in different domains for the three datasets for cross-scene classification.

C. Results and Analysis

To perform the domain generalization ability of the proposed model, we conduct a series of experiments on the UDAD to

(a)

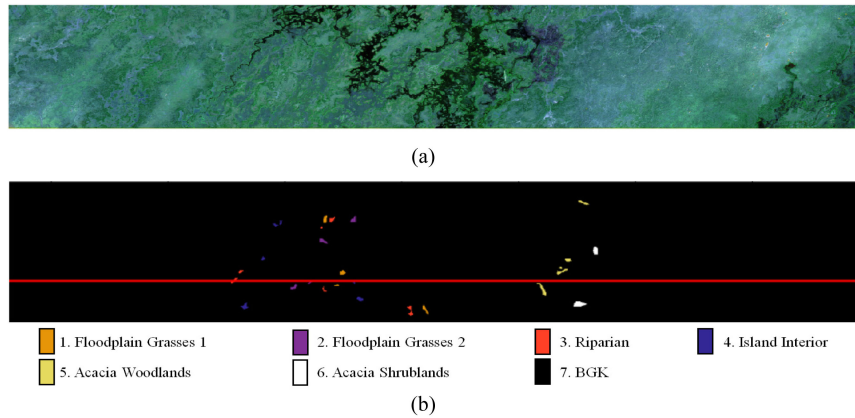| 1. Floodplain Grasses 1 | 2. Floodplain Grasses 2 | 3. Riparian | 4. Island Interior |
| 5. Acacia Woodlands | 6. Acacia Shrublands | 7. BGK | |

(b)

Fig. 6. Botswana scene. (a) False-color image with band (15, 37, 86). (b) Colorful ground-truth image. The upper and lower parts of the red line are selected as the source domain and target domain, respectively.

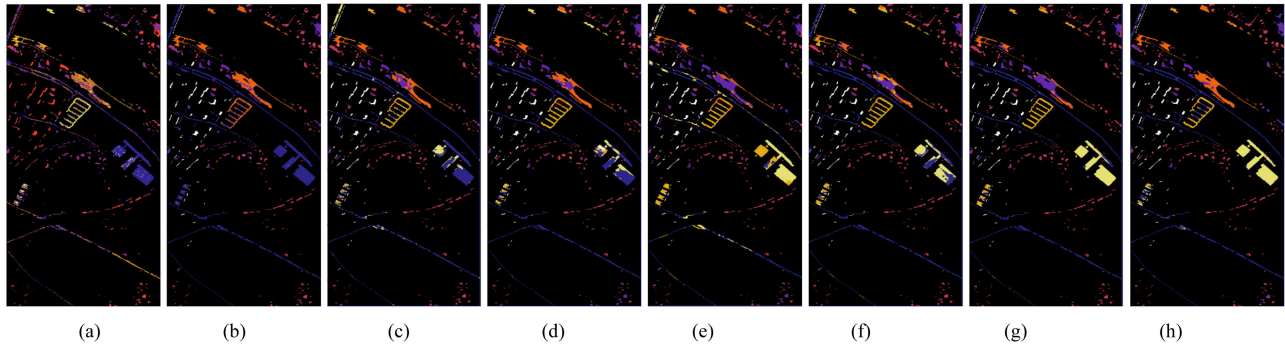| (a) | (b) | (c) | (d) | (e) | (f) | (g) | (h) |

Fig. 7. Classification results of the Pavia Center with compared methods. (a) KNN. (b) OSDA. (c) DANN. (d) UDHS. (e) HS-ADDA. (f) HS-D1. (g) HS-D2. (h) UDAD.

TABLE III
NUMBER OF TRAINING AND CLASSIFICATION MODEL FOR THE HSI DATASETS

Phase	Pavia University →Pavia Center	Pavia Center →Pavia University	Botswana (Top scene→ Bottom scene)
Source Training	100 (labeled)	100 (labeled)	20 (labeled)
Target Training	100 (unlabeled)	100 (unlabeled)	20 (unlabeled)
Target Classification	Unsupervised	Unsupervised	Unsupervised

tackle the HSIC with different domain shifts. First, we train the Pavia University scene as the source domain and perform an unsupervised classification on the Pavia Center image. In this experiment, we randomly select 100 samples for each class in the source HSI to pretrain the dense CNN. The colorful classification maps and the unsupervised classification results are shown in Fig. 7 and Table IV. As given in Table V, the proposed model generates the best OA of 87.97%, AA of 88.02%, and Kappa of 85.50%, which is better than other transfer learning methods. Also, the HS-D1 generates the OA of 73.36%, and the HS-D2 obtains the OA of 81.49%, which illustrates the feasibility of the compact structure for the feature extraction. It is noted that the KNN with ten primary components yields the worst performance. We also conducted the experiments with the number changed from 5 to 25 with an interval of five to evaluate

the effect of the different numbers of the primary components. Consequently, the values of OA are 34.18% and 32.25% with numbers 5 and 25, respectively. And the values of AA vary from 37.74% to 36.25%. The obtained results illustrate that the component number has less effect on the KNN method. The UDHS approach implemented the unsupervised classification with the classifiers trained in the source domain. As we can observe, the OA is provided of 68.28% with UDHS, which means the designed feature compaction supplies better representation than the approach without domain adaptation in this situation.

The second experiment is tested on the unlabeled Pavia University with the labeled Pavia Center as the source domain. We adopt 100 source samples for each class in the Pavia Center. The same parameters are set same as the first experiment. Fig. 8 and Table V demonstrate the classification results of the related models. The UDHS method yields the worst performance of HS-ADDA with 73.08%, and it shows unsatisfactory results without dense-based compaction. Compared with other methods, our proposed method generates the best performance than other compared methods. Moreover, the proposed UDAD achieves better accuracy in most classes. The Pavia University and Pavia Center have more domain discrepancies in the two scenes. The comparison results reflect our method achieves a favorable effect on the cross scenes with more differences of the domains.

TABLE IV
CLASSIFICATION PERFORMANCE FROM THE RESULTS WITH ALL THE COMPARED METHODS OF THE PAVIA CENTER (PAVIA UNIVERSITY→PAVIA CENTER)

Class P_{OA}%	KNN	OSDA	DANN	UDHS	HS-ADDA	HS-D1	HS-D2	UDAD (Our)
1	36.51	77.90±16.63	89.09±3.65	86.67±12.60	82.59±12.12	92.34±3.01	87.87±12.92	**94.24±4.37**
2	58.42	67.97±19.16	79.40±8.01	77.23±20.72	**86.98±4.97**	62.81±22.35	70.87±10.08	71.18±4.34
3	37.37	94.02±3.95	95.92±2.92	89.36±6.52	71.03±11.61	**94.10±6.14**	92.38±8.63	87.70±6.59
4	53.23	**85.77±7.57**	78.35±15.55	71.19±18.06	85.96±11.30	59.69±26.10	78.06±17.10	84.58±3.20
5	73.64	6.78±3.68	84.77±29.14	65.11±31.83	25.65±21.22	87.31±17.08	**90.05±7.52**	85.24±4.41
6	2.44	32.57±30.09	36.47±34.30	18.46±19.01	24.92±25.77	40.16±30.81	68.22±19.21	**93.61±8.00**
7	0	**100.00±0.00**	99.93±9.79	99.61±0.44	90.10±15.68	97.88±3.68	86.99±15.77	99.58±0.25
OA	34.54	69.16±3.22	76.98±4.77	68.28±1.13	66.03±6.15	73.36±4.82	81.49±6.04	**87.97±0.60**
AA	37.37	66.43±1.61	80.56±4.00	72.16±2.92	66.37±6.65	76.33±5.32	82.06±4.36	**88.02±0.28**
Kappa	22.95	63.07±3.52	72.24±5.83	62.01±1.02	59.50±7.15	68.10±5.84	77.76±7.13	**85.50±0.70**
Time (s)	4.83	74.24	539.03	157.89	494.24	629.58	771.07	2337.60

The boldface number indicates the maximum value of each row of the comparison.

TABLE V
CLASSIFICATION PERFORMANCE FROM THE RESULTS WITH ALL THE COMPARED METHODS OF THE PAVIA UNIVERSITY (PAVIA CENTER→PAVIA UNIVERSITY)

Class P_{OA}%	KNN	OSDA	DANN	UDHS	HS-ADDA	HS-D1	HS-D2	UDAD (Our)
1	4.22	24.73±24.46	73.84±8.85	92.19±3.44	**97.73±2.91**	82.77±5.87	76.81±8.40	79.16±6.49
2	12.5	76.11±21.97	62.75±5.07	39.64±8.18	57.64±5.17	63.53±8.65	64.27±10.09	**79.20±6.84**
3	91.71	47.55±27.69	92.32±3.88	**95.51±3.32**	83.98±7.21	88.72±4.17	91.52±2.41	92.72±2.51
4	62.82	24.77±14.64	61.33±3.56	58.27±8.88	59.90±13.70	61.84±10.13	57.61±18.10	47.85±7.22
5	70.72	13.05±22.58	71.79±24.92	35.50±27.78	7.75±5.96	80.68±22.29	87.73±10.47	**91.95±4.05**
6	91.88	22.54±29.07	**91.36±4.16**	34.97±16.37	6.76±6.44	40.84±26.52	56.75±35.70	46.21±15.86
7	99.05	90.71±10.04	**95.79±2.00**	92.43±4.39	75.21±7.25	74.21±1.07	74.95±2.03	89.91±7.92
OA	33.93	51.86±6.92	69.40±2.72	55.87±3.56	60.78±1.56	69.66±4.64	69.89±2.28	**76.56±1.97**
AA	61.84	42.78±8.95	**78.46±4.09**	64.06±4.20	55.71±1.49	70.37±4.02	72.81±4.69	75.29±2.00
Kappa	24.91	32.61±9.86	60.01±3.59	45.80±3.19	48.83±1.18	60.34±5.16	60.66±2.31	**67.82±2.16**
Time (s)	4.88	74.21	655.20	174.34	930.52	616.36	756.36	1988.76

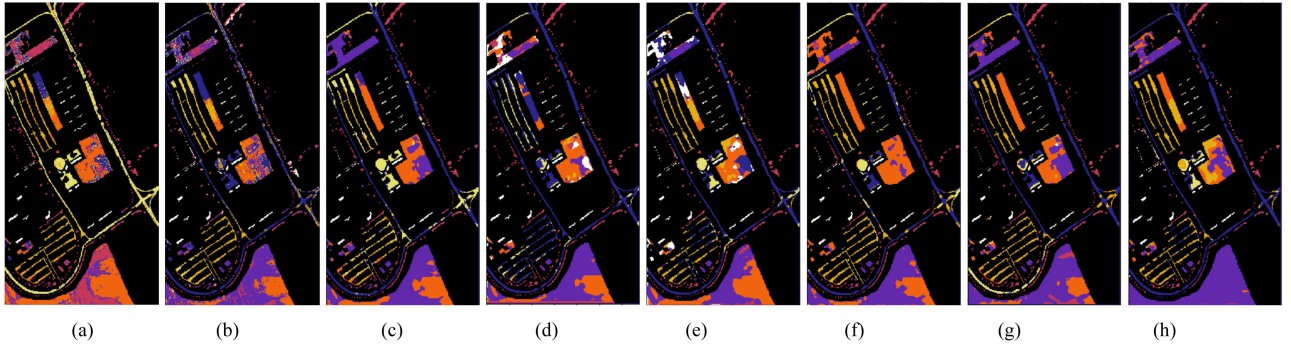The boldface number indicates the maximum value of each row of the comparison.

Fig. 8. Classification results of the Pavia University with compared methods. (a) KNN. (b) OSDA. (c) DANN. (d) UDHS. (e) HS-ADDA. (f) HS-D1. (g) HS-D2. (h) UDAD.

The last experiment is conducted in Botswana, the source domain is the top half of the dataset, and the target domain is the bottom half. The cross domains have the same class category, which means the few domain discrepancies presented in the HSIC. We choose six classes with 20 samples per class for the following experiment. The classification results are given in Table VI and Fig. 9. We observe that the KNN-related approaches have the lowest OA and Kappa in this experiment. Our method achieves the best performance in terms of the three evaluation indexes. The OA value reaches 96.09%, the AA is 96.45%, and the Kappa is the highest of 95.24%. In particular, the UDAD succeeds the best accuracy for the most classes in this situation. The DANN approach generated the OA of 92.89%,

and the HS-ADDA yields the value of 60.78%. It is noted that the mentioned DANN and HS-ADDA are implemented by the adversarial scheme without dense-based feature compaction, which demonstrated that our approach is beneficial for the domain discrepancy reduction and provided with better accuracy. Moreover, we perform the proposed method on an unsupervised domain adaptation with the source and target domains exchanged of the Botswana data. Specifically, the setting of the experiment is the same as the situation with the previous scene. As expected, we acquire the OA of 93.76%, the AA of 93.55%, and the Kappa of 92.36%. Clearly, the source domain of top Botswana contributes a more promising adaptation for reducing domain discrepancy than the source of the bottom scene,

TABLE VI
CLASSIFICATION PERFORMANCE FROM THE RESULTS WITH ALL THE COMPARED METHODS OF BOTSWANA (TOP SCENE→BOTTOM SCENE)

Class $P_{OA}\%$	KNN	OSDA	DANN	UDHS	HS-ADDA	HS-D1	HS-D2	UDAD (Our)
1	68.09	83.80±8.68	96.69±2.18	66.13±4.72	63.31±4.78	81.47±10.86	89.57±14.53	**100±0.00**
2	6.33	100.00±0.00	94.57±4.36	97.72±2.08	94.94±22.97	**100.00±0.00**	99.75±0.51	**100±0.00**
3	23.29	71.37±14.24	89.19±7.78	27.81±2.20	25.48±2.96	88.90±10.29	89.18±6.18	**89.21±8.88**
4	0	81.77±5.36	**100.00±0.00**	78.44±8.45	71.43±11.06	**100.00±0.00**	98.18±3.64	**100±0.00**
5	19.23	29.42±18.62	75.87±20.82	**100.00±0.00**	**100.00±0.00**	86.92±17.47	91.92±5.36	87.98±9.96
6	0.65	83.51±2.19	99.12±1.46	**100.00±0.00**	**100.00±0.00**	99.61±0.78	99.87±0.26	**100.00±0.00**
OA	23.65	74.97±3.10	92.89±4.07	75.24±1.75	73.08±2.10	91.62±2.44	94.05±3.32	**96.09±0.80**
AA	19.59	74.99±4.34	92.57±4.55	78.35±1.96	75.86±2.75	92.82±2.39	94.74±2.35	**96.45±0.57**
Kappa	8.01	69.38±4.32	91.25±5.02	70.08±2.13	67.43±2.61	89.83±2.95	92.79±3.98	**95.24±0.97**
Time (s)	2.57	20.35	157.50	43.45	79.57	208.52	256.34	302.20

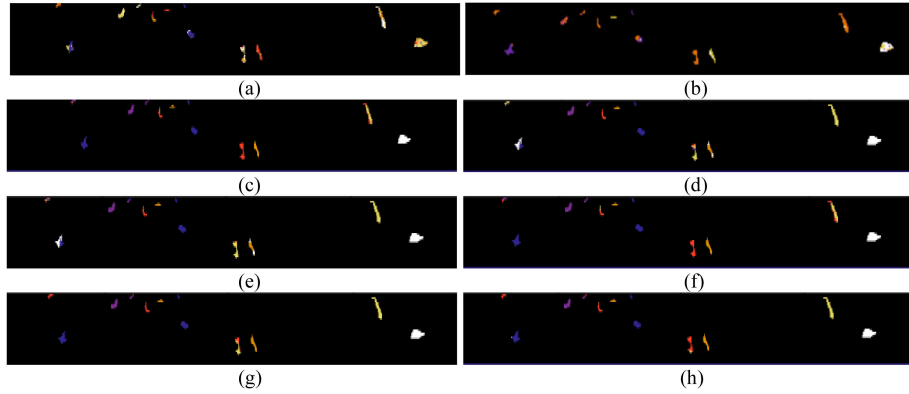The boldface number indicates the maximum value of each row of the comparison.

Fig. 9. Classification results of the Botswana data with compared methods. (a) KNN. (b) OSDA. (c) DANN. (d) UDHS. (e) HS-ADDA. (f) HS-D1. (g) HS-D2. (h) UDAD.
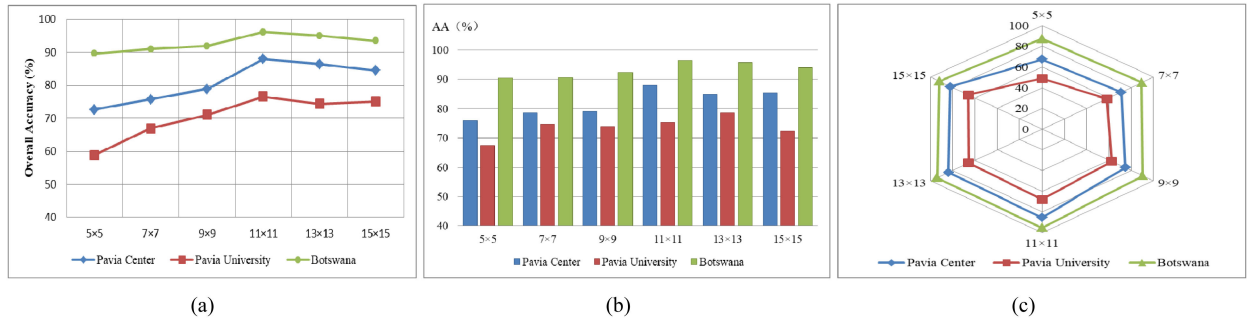

Fig. 10. Effects of the HSIC performance with the different patch sizes for the four HSI scenes with the UDAD. (a) OA. (b) AA. (c) Kappa. (average of five runs).

which generates better performance on the unlabeled target domain.

For the training time of the three cross-scenes, the proposed method costs more than other methods due to the addition of the adversarial training time. The HS-D1 and HS-D2 are more efficient than the UDAD, showing that the proposed method with the two dense blocks is more time-consuming and significantly affects feature adaptation. The KNN approach costs less time since no training procedure in the unsupervised HSIC implementation generates the worst performances for the three datasets.

In conclusion, compared with other modern techniques, the UDAD model is adequate for cross-scene transfer learning tasks. Besides, in contrast to the exhaustive fine-tuning process, the proposed model implements the spectral–spatial feature adaptation with the adversarial mechanism to align the feature distribution of various domains.

D. Analysis of the Parameter Effect

In this section, we analyze the influence of two parameters on the performances of the proposed model, including image patch size and the number of the selected source sample. The classification performances of the patch size towards the three scenes are shown in Fig. 10. As shown in Fig. 10(a), the vertical axis represents the classification accuracy, and we observed that
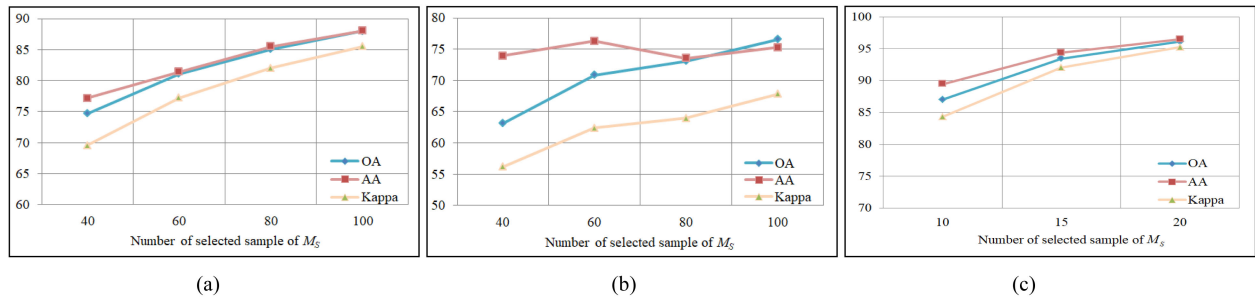
Fig. 11. Classification performance obtained with the increase of the sample number from the source domain. (a) Pavia Center. (b) Pavia University. (c) Botswana. (average of five runs).
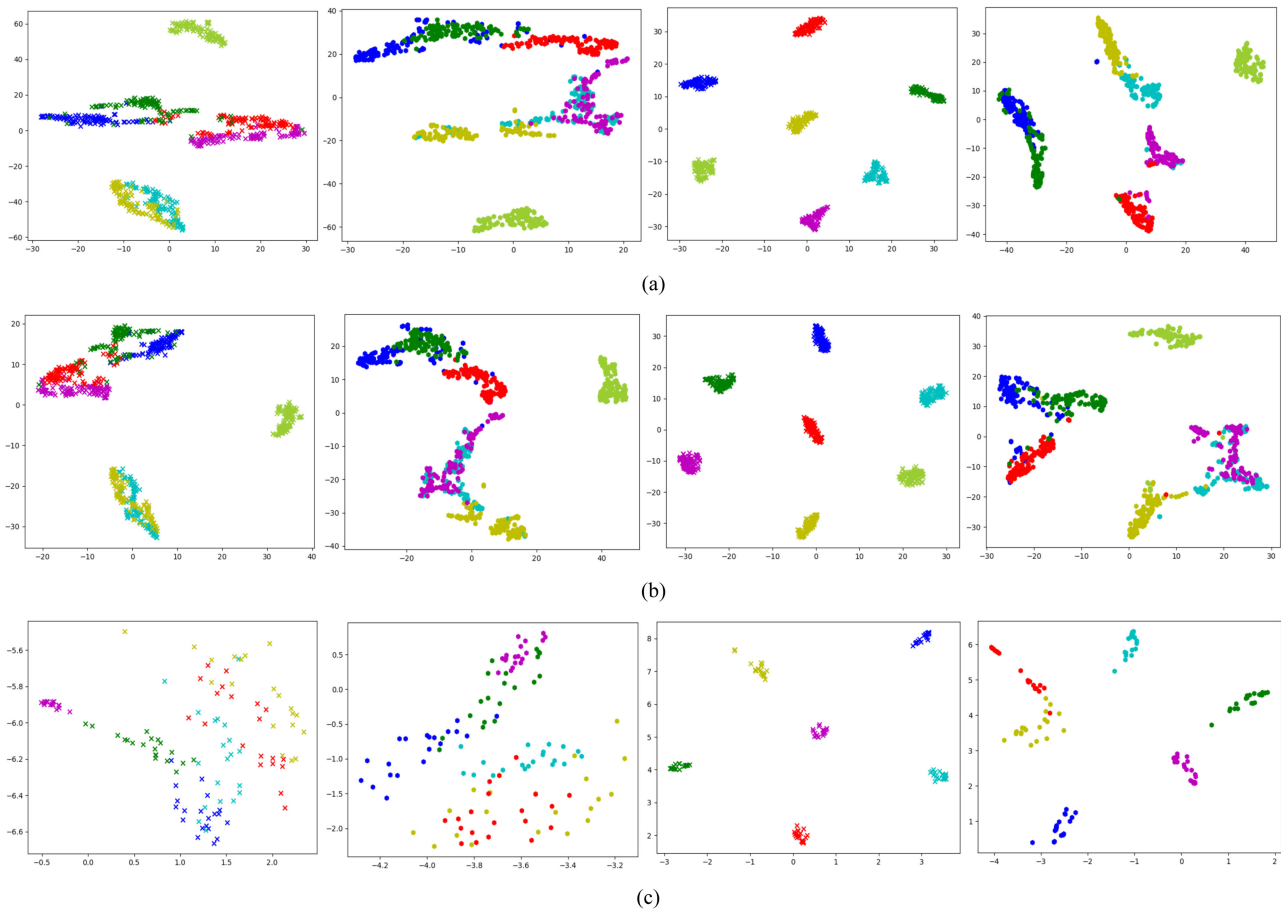

Fig. 12. t-SNE visualization of the feature maps of different domains. The left and the second column show the original distributions of the source and target HSI. The third and the right column illustrate the feature representation extracted with the dense CNN module of the source and target scenes, respectively. The symbol "x" represents the source domain and "o" represents the target domain. The same colors mean the same classes of scenes. (a) Pavia Center (seven classes and 100 samples per class). (b) Pavia University (seven classes and 100 samples per class). (c) Botswana (six classes and 20 samples per class).

the patch size is crucial for the OA value. The highest OAs are yielded with the size of 11×11 for the Pavia Center (87.97%), the Pavia University (76.56%), and Botswana (96.09%). When the patch size is set to 5×5, the worst performances are generated by 72.66%, 58.88%, and 89.57% for the Pavia Center, the Pavia University, and Botswana. The histograms in Fig. 10(b) reflect the AA with different patch sizes. It can be observed that the proposed model achieves the highest AA with the size of 11×11 for the Pavia Center (88.02%) and Botswana (96.45%).

The Pavia University obtains the best AA of 78.62% with a size of 13×13. The relationship between the Kappa with different patch sizes is shown in Fig. 10(c). Specifically, the Kappa performs the highest value with the size of 11×11 for the three scenes, which are 85.5%, 67.82%, and 95.24%, respectively. The effect of the number of the source domain for pretraining is shown in Fig. 11. We compare the classification performance with the sample number of $\{40, 60, 80, 100\}$ for the Pavia datasets. As shown in the polylines of Fig. 11(a) and (b),
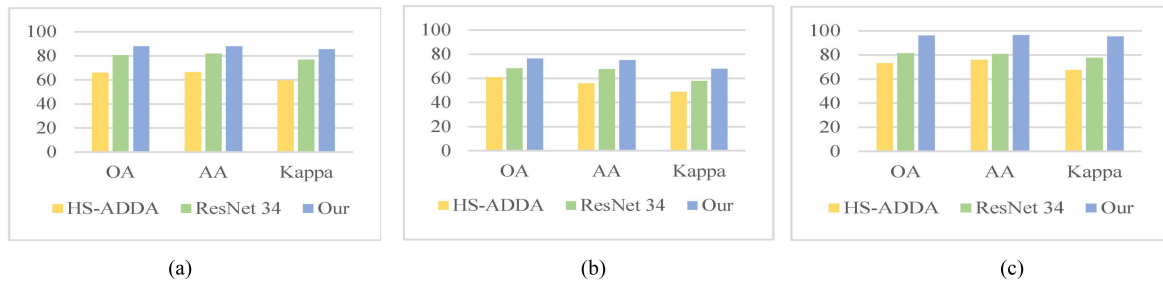
Fig. 13. Performance of three cross-scenes of HSIC with different backbone networks. (a) Pavia University→Pavia Center. (b) Pavia Center→Pavia University. (c) Top Botswana→Bottom Botswana.

we see that the better OAs are obtained with more numbers of the training sample. For Botswana, we conducted the experiments with the sample number of {10, 15, 20}. The line chart of Botswana is shown in Fig. 11(c). Similarly, the more training sample from the source domain generates a better performance of OA, AA, and Kappa. In summary, we conclude that both the patch size and the sample number affect the capability of the domain adaptation.

E. Analysis of the Backbone Network

In this section, we perform experiments to verify the performance of the backbone network of the feature extraction module. Briefly, the compared networks consist of the traditional CNN network with two convolution layers that is recorded as the HS-ADDA approach, the ResNet 34, and our dense-based network. Intuitively, as shown in Fig. 13, our implementation with dense-based compaction explicitly implies abundant information according to the labels of the source domain and is beneficial to the feature alignment of the adaptation phase. Especially, the OA value reaches 96.09% for the Botswana scene with our approach. Notably, the performance variation with the other two networks indicates that ResNet 34 network obtains a higher value of accuracy than the traditional CNN network. Due to the complex character of the hyperspectral scene, the backbone based on the traditiaonal CNN generated unsatisfactory feature refinement, which brings the OA of 66.03%, 60.78%, and 73.08% for the Pavia University→Pavia Center, Pavia Center→Pavia University, and Top Botswana→Bottom Botswana, respectively.

F. Visualization of Feature Adaptation

To simulate feature distributions extracted from the different domains with the proposed UDAD model, Fig. 12 presents the feature visualization of the three HSI datasets by the t-SNE algorithm. It can be observed that the visualization of the features before the domain adaption is disorganized in the left two columns. Apparently, the initial distributions in the two different spaces are quite different before adversarial adaptation. In contrast, the source and target samples are discriminated after the feature extraction CNN, and the pixels from the same classes are clustered to separate effortlessly. Specifically, Botswana clusters are more separable than the other two datasets, and the

Pavia Center tends to show a better clustering result than the Pavia University after the adversarial transfer learning.

IV. CONCLUSION

This work addressed the HSIC cross-scene classification through the adversarial adaption of different domains. We presented an unsupervised domain adaptation classification architecture for unsupervised HSIC. The spectral–spatial feature alignment was implemented with a generative and adversarial framework, which was the essential component of the proposed architecture. Besides, the designed dense CNN network as a backbone was employed to capture the abundant discriminative feature of the domains. Experimental results and analysis demonstrate that the proposed UDAD method was applicable to different domain discrepancies of the HSI scenes and validates the robust performance on the hyperspectral unsupervised classification.

REFERENCES

[1] P. Ghamisi *et al.*, "Advances in hyperspectral image and signal processing: A comprehensive overview of the state of the art," *IEEE Geosci. Remote Sens. Mag.*, vol. 5, no. 4, pp. 37–78, Dec. 2017.

[2] L. Zhang, L. Zhang, and B. Du, "Deep learning for remote sensing data: A technical tutorial on the state of the art," *IEEE Geosci. Remote Sens. Mag.*, vol. 4, no. 2, pp. 22–40, Jun. 2016.

[3] C. Li, Y. Ma, X. Mei, C. Liu, and J. Ma, "Hyperspectral image classification with robust sparse representation," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 5, pp. 641–645, May 2016.

[4] Z. Yan *et al.*, "HD-CNN: Hierarchical deep convolutional neural network for large scale visual recognition," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2015, pp. 2740–2748.

[5] X. Ma, A. Fu, J. Wang, H. Wang, and B. Yin, "Hyperspectral image classification based on deep deconvolution network with skip architecture," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 8, pp. 4781–4791, Aug. 2018.

[6] L. Fang, C. Wang, S. Li, and J. A. Benediktsson, "Hyperspectral image classification via multiple-feature-based adaptive sparse representation," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 7, pp. 1646–1657, Jul. 2017.

[7] K. Zhang, W. Zuo, and L. Zhang, "FFDNet: Toward a fast and flexible solution for CNN-Based image denoising," *IEEE Trans. Image Process.*, vol. 27, no. 9, pp. 4608–4622, Sep. 2018.

[8] J. M. Haut, M. E. Paoletti, J. Plaza, A. Plaza, and J. Li, "Visual attention-driven hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 10, pp. 8065–8080, Oct. 2019.

[9] L. Zhang, L. Zhang, D. Tao, X. Huang, and B. Du, "Hyperspectral remote sensing image subpixel target detection based on supervised metric learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 8, pp. 4955–4965, Aug. 2014.

[10] C. Yu, Y. Wang, M. Song, and C. Chang, "Class signature-constrained background- suppressed approach to band selection for classification of hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 1, pp. 14–31, Jan. 2019.

[11] Q. Du, "Modified Fisher's linear discriminant analysis for hyperspectral imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 4, no. 4, pp. 503–507, Oct. 2007.

[12] C. Chang, "Statistical detection theory approach to hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 4, pp. 2057–2074, Apr. 2019.

[13] H. Yu, L. Gao, W. Li, Q. Du, and B. Zhang, "Locality sensitive discriminant analysis for group sparse representation-based hyperspectral imagery classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 8, pp. 1358–1362, Aug. 2017.

[14] X. Kang, S. Li, and J. A. Benediktsson, "Spectral–spatial hyperspectral image classification with edge-preserving filtering," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 5, pp. 2666–2677, May 2014.

[15] C. Yu *et al.*, "Hyperspectral image classification method based on CNN architecture embedding with hashing semantic feature," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 6, pp. 1866–1881, Jun. 2019.

[16] X. Ma, H. Wang, and J. Geng, "Spectral–spatial classification of hyperspectral image based on deep auto-encoder," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 9, pp. 4073–4085, Sep. 2016.

[17] S. Wan, C. Gong, P. Zhong, B. Du, L. Zhang, and J. Yang, "Multiscale dynamic graph convolutional network for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 5, pp. 3162–3177, May 2020.

[18] D. Hong, L. Gao, J. Yao, B. Zhang, A. Plaza, and J. Chanussot, "Graph convolutional networks for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 7, pp. 5966–5978, Aug. 2020.

[19] L. Zhu, Y. Chen, P. Ghamisi, and J. A. Benediktsson, "Generative adversarial networks for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 9, pp. 5046–5063, Sep. 2018.

[20] H. Wang, C. Tao, J. Qi, H. Li, and Y. Tang, "Semi-supervised variational generative adversarial networks for hyperspectral image classification," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2019, pp. 9792–9794.

[21] S. Pande, A. Banerjee, S. Kumar, B. Banerjee, and S. Chaudhuri, "An adversarial approach to discriminative modality distillation for remote sensing image classification," in *Proc. Int. Conf. Comput. Vis.*, 2019, pp. 4571–4580.

[22] X. Lu, X. Zheng, and Y. Yuan, "Remote sensing scene classification by unsupervised representation learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 9, pp. 5148–5157, Sep. 2017.

[23] W. Song, S. Li, L. Fang, and T. Lu, "Hyperspectral image classification with deep feature fusion network," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 6, pp. 3173–3184, Jun. 2018.

[24] Y. Chen, K. Zhu, L. Zhu, X. He, P. Ghamisi, and J. A. Benediktsson, "Automatic design of convolutional neural network for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 9, pp. 7048–7066, Sep. 2019.

[25] Y. Li, W. Xie, and H. Li, "Hyperspectral image reconstruction by deep convolutional neural network for classification," *Pattern Recognit.*, vol. 63, pp. 371–383, 2017.

[26] W. Hu, Y. Huang, L. Wei, F. Zhang, and H. Li, "Deep convolutional neural networks for hyperspectral image classification," *J. Sensors*, vol. 2015, pp. 1–12, Jan. 2015.

[27] S. Mei, J. Ji, Y. Geng, Z. Zhang, X. Li, and Q. Du, "Unsupervised spatial–spectral feature learning by 3D convolutional autoencoder for hyperspectral classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 9, pp. 6808–6820, Sep. 2019.

[28] S. K. Roy, G. Krishna, S. R. Dubey, and B. B. Chaudhuri, "HybridSN: Exploring 3-D–2-D CNN feature hierarchy for hyperspectral image classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 2, pp. 277–281, Feb. 2020.

[29] R. R. Hang, Q. Liu, D. Hong, and P. Ghamisi, "Cascaded recurrent neural networks for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 8, pp. 5384–5394, Aug. 2019.

[30] J. Shen, X. Cao, Y. Li, and D. Xu, "Feature adaptation and augmentation for cross-scene hyperspectral image classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 4, pp. 622–626, Apr. 2018.

[31] M. Ye, Y. Qian, J. Zhou, and Y. Y. Tang, "Dictionary learning-based feature-level domain adaptation for cross-scene hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 3, pp. 1544–1562, Mar. 2017.

[32] D. Tuia, E. Pasolli, and W. J. Emery, "Using active learning to adapt remote sensing image classifiers," *Remote Sens. Environ.*, vol. 115, no. 9, pp. 2232–2242, Sep. 2011.

[33] J. Peng, W. Sun, L. Ma, and Q. Du, "Discriminative transfer joint matching for domain adaptation in hyperspectral image classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 6, pp. 972–976, Jun. 2019.

[34] J. Lin, C. He, Z. J. Wang, and S. Li, "Structure preserving transfer learning for unsupervised hyperspectral image classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 10, pp. 1656–1660, Oct. 2017.

[35] J. Yang, Y. Zhao, and J. C. Chan, "Learning and transferring deep joint spectral–spatial features for hyperspectral classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 8, pp. 4729–4742, Aug. 2017.

[36] Y. Jiang, Y. Li, and H. Zhang, "Hyperspectral image classification based on 3-D separable resnet and transfer learning," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 12, pp. 1949–1953, Dec. 2019.

[37] C. Zhong, J. Zhang, S. Wu, and Y. Zhang, "Cross-scene deep transfer learning with spectral feature adaptation for hyperspectral image classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 2861–2873, 2020.

[38] I. Goodfellow, J. Abadie, M. Mirza, Xu B, and Y. Bengio, "Generative adversarial nets," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 27, no. 3, 2014, pp. 2672–2680.

[39] E. Tzeng, J. Hoffman, K. Saenko, and T. Darrell, "Adversarial discriminative domain adaptation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 2962–2971.

[40] L. Song *et al.*, "Unsupervised domain adaptive re-identification: Theory and practice," *Pattern Recognit.*, vol. 102, no. 107173, pp. 1–11, Jun. 2020.

[41] L. Zhang, M. Lan, J. Zhang, and D. Tao, "Stagewise unsupervised domain adaptation with adversarial self-training for road segmentation of remote-sensing images," in *IEEE Trans. Geosci. Remote Sens.*, early access, Aug. 12, 2021, doi: 10.1109/TGRS.2021.3104032.

[42] Y. Ganin *et al.*, "Domain adversarial training of neural networks, " *J. Mach. Learn. Res.*, vol. 17, no. 59, pp. 1–35, 2016.

[43] V. S. Nirmal and K. P. Soman, "Open set domain adaptation for hyperspectral image classification using generative adversarial network," in *Inventive Commun. Comput. Technol.*, 2020, pp. 819–827

Chunyan Yu (Member, IEEE) received the B.Sc. and Ph.D. degrees in environment engineering from Dalian Maritime University, Dalian, China, in 2004 and 2012, respectively.

She is currently an Associate Professor with the Information Science and Technology College, Dalian Maritime University. Her research interests include image segmentation, hyperspectral image classification, and pattern recognition.

Caiyu Liu (Member, IEEE) received the B.E. degree in network engineering from the Zhongyuan University of Technology, Zhengzhou, China, in 2019. She is currently working toward the M.A. degree in computer science and technology with the Dalian Maritime University, Dalian, China.

Her research interests include hyperspectral image processing and deep learning.

Haoyang Yu (Member, IEEE) received the B.S. degree in information and computing science from Northeastern University, Shenyang, China, in 2013 and the Ph.D. degree in cartography and geographic information system from the Key Laboratory of Digital Earth Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, China, 2019.

He is currently a Xing Hai Associate Professor with the Center of Hyperspectral Imaging in Remote Sensing, Information Science and Technology College, Dalian Maritime University, Dalian, China. His research interests include models and algorithms for hyperspectral image processing, analysis, and applications.

Meiping Song (Member, IEEE) received the Ph.D. degree in the Computer Science and Technology, Harbin Engineering University, Harbin, China, in 2006.

From 2013 to 2014, she was a Visiting Associate Research Scholar with the University of Maryland, Baltimore County. She is currently an Associate Professor with the College of Information Science and Technology, Dalian Maritime University, Dalian, China. Her research interests include remote sensing and hyperspectral image processing.

Chein-I Chang (Life Fellow, IEEE) received the B.S. degree in mathematics from Soochow University, Taipei, Taiwan, the M.S. degree in mathematics from the Institute of Mathematics, National Tsing Hua University, Hsinchu, Taiwan, the M.A. degree in mathematics from the State University of New York at Stony Brook, Stony Brook, NY, USA, in 1977, the M.S. degree in theoretical computation science and M.S.E.E. degree in electrical engineering from the University of Illinois at Urbana–Champaign, Champaign, IL, USA, in 1980 and 1982, and the Ph.D. degree in electrical engineering from the University of Maryland, College Park, MD, USA, in 1987.

Since 1987, he has been with the University of Maryland, Baltimore County, where he is currently a Professor with the Department of Computer Science and Electrical Engineering. He is also currently holding the Chang Jiang Scholar Chair Professorship and has been the director of Center for Hyperspectral Imaging in Remote Sensing, Dalian Maritime University, Dalian, China, since 2016, Hua Shan Scholar Chair Professorship with Xidian University, Xian, China, since 2016, and Feng-Tai Chair Professor with National Yunlin University of Science and Technology, since 2017. His research interests include remote sensing and hyperspectral image processing.