Generalized Zero-Shot Domain Adaptation for Unsupervised Cross-Domain PolSAR Image Classification

Rong Gui, Xin Xu, Rui Yang, Kailiang Deng, and Jun Hu, *Member, IEEE*

***Abstract*—Cross-domain polarimetric synthetic aperture radar interpretation is urgently needed, due to the rapid data growth and label scarcity. However, the class distribution shift problems hinder the reuse of labeled samples among cross-domain images. Most of the existing domain adaptations can only handle the cross-domain case of same categories between source and target domains, while the categories of the target domain are usually more abundant than those of the source domain. To improve the usability of labeled samples among cross-domain images, an unsupervised generalized zero-shot domain adaptation (uGZSDA) based on scattering component semantics (SCSs) is proposed. By using SCSs and limited labeled samples (seen categories) in the source domain, more land cover types (seen and unseen categories) in the unlabeled target domain can be inferred. First, a stacked autoencoder (SAE) extracts source/target-domain features, and SCSs of typical land covers are constructed by cross-domain databases and statistical scattering components. Second, combining SAE features and source-domain samples, the most likely seen class samples in the target domain are selected by probability sorting, and the SAE is retrained by obtained selected seen samples. Third, the unseen class samples in the target domain are inferred by the retrained SAE, classification probability, and semantic similarity. Finally, the selected seen and inferred unseen class samples in the target domain are used to further retrain the SAE, and the target domain is classified by the retrained SAE and the classifier. The proposed uGZSDA is verified among 16 cross-domain PolSAR datasets. Using SCS and two to three types of seen samples from the source domain, the accuracies of seven types of land covers in the unlabeled target domain can reach 76–83.96%.**

***Index Terms*—Generalized zero-shot learning (GZSL), polarimetric synthetic aperture radar (PolSAR) image, scattering component semantics (SCS), stacked autoencoder (SAE), unsupervised domain adaptation (uDA).**

Manuscript received August 3, 2021; revised October 15, 2021 and November 24, 2021; accepted December 8, 2021. Date of publication December 13, 2021; date of current version December 29, 2021. This work was supported in part by the National Natural Science Foundation of China under Grant 42030112 and in part by the Nature Science Foundation of Hunan Province under Grant 2020JJ2043. *(Corresponding authors: Xin Xu; Kailiang Deng.)*

Rong Gui and Jun Hu are with the School of Geosciences and Info-Physics, Central South University, Changsha 410083, China (e-mail: ronggui2013@whu.edu.cn; csuhujun@csu.edu.cn).

Xin Xu and Rui Yang are with the School of Electronic Information, Wuhan University, Wuhan 430072, China (e-mail: xinxu@whu.edu.cn; ruiyang@whu.edu.cn).

Kailiang Deng is with the Naval Institute of Hydrographic Surveying and Charting, Tianjin 300061, China (e-mail: dkl_hych@163.com).

Digital Object Identifier 10.1109/JSTARS.2021.3134766

I. Introduction

POLARIMETRIC synthetic aperture radar (PolSAR) can provide rich polarimetric scattering information for observed land covers under full-time and all-weather conditions [1], [2]. With the rapid growth of PolSAR data, it is particularly urgent to maximize the role of limited labeled samples in cross-domain (cross-dataset, cross-temporal, and cross-source) PolSAR interpretation [3]–[5]. Specifically, an increasing number of spaceborne, airborne, and UAV-borne synthetic aperture radar (SAR) sensors, e.g., Radarsat-2, Gaofen-3, and UAVSAR, provide abundant PolSAR data. However, accurately labeled samples are usually rare, and labeling the samples always needs professional knowledge in PolSAR interpretation [6], [7]. In addition, due to the differences of sensors, imaging geometries, and land cover distributions, the class distribution shift (CDS) problems [5], [8] between cross-domain data hinder the large-scale machine-learning-based applications in PolSAR images. It is acknowledged that the traditional supervised learning and deep learning methods can hardly deal with the cross-domain data with CDS [3]–[5]. Cross-domain task mainly includes cross-dataset (cross-region) [12], cross-temporal [13], and cross-source [14]. Generally, as long as there are imaging parameter differences or scene differences between the source domain and the target domain, they can be defined as cross-domain problems [3], [9]. Therefore, exploiting and reusing existing labeled samples to classify newly acquired data is one of the most concerned hotspots and difficulties in PolSAR interpretation [3], [9].

As a particular case of transfer learning, an unsupervised domain adaptation (uDA) can transfer supervised information of the existing source domain to those of the unlabeled target domain by reducing the CDS problems [5], [10], [11]. The uDA is increasingly concerned and applied to solve the label scarcity problem in remote sensing analysis [15]–[18]. However, most of the aforementioned uDAs only focus on those cases where source and target domains are with exactly the same land covers [19], as shown at the top of Fig. 1. In many practical remote sensing applications, there are usually unseen or unknown categories in the newly acquired target-domain data [20], such as the cases at the bottom of Fig. 1. For the unlabeled target data, it is necessary to reuse the existing labeled samples (seen classes) from the source domain to improve the interpretation accuracy and identify the unseen classes that have not been

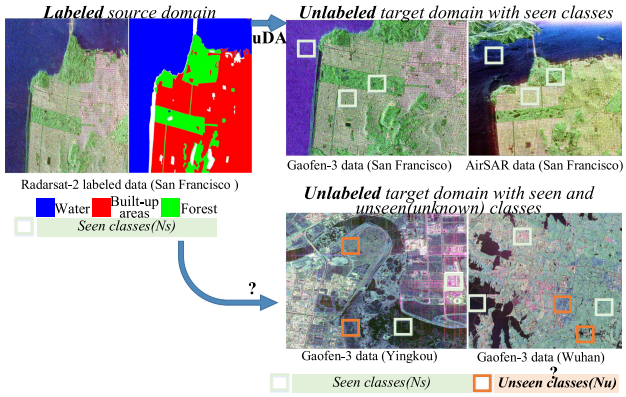


Fig. 1. Illustration of typical uDA and seen and unseen classes in cross-domain PolSAR classification scenarios.

trained and labeled in the source domain. In this case, most of the existing domain adaptations cannot meet this practical application requirement [19], [20].

Recently, zero-shot domain adaptation (ZSDA) and generalized zero-shot domain adaptation (GZSDA) have been developed and concerned [21]–[23] in the computer vision field as promising solutions to the above problems. They have demonstrated the ability to infer and identify unseen categories in the target domain by using the seen samples and semantic information [24]–[28]. However, both ZSDA and GZSDA need labeled seen samples from source and target domains. Recently, these domain adaptations that can recognize unseen classes have been concerned in the optical remote sensing field [19].

Many of the existing cross-domain methods designed for natural images and optical remote sensing data cannot be directly applied to PolSAR interpretation [5], [9]. Moreover, it is usually more common for the newly obtained target domain to not contain any annotations in PolSAR applications. The difficulties of cross-domain PolSAR interpretation mainly come from three aspects.

1) More imaging factors affect and exacerbate the distribution shifts between cross-domain data, including sensors, incidence angles, ascending and descending orbits, operating bands, speckle, imaging seasons, and so on [3], [29].
2) For uDA and unsupervised generalized zero-shot learning domain adaptation (uGZSDA) problems, feature expressions should be applicable for any unlabeled target domain, i.e., the cross-domain features should overcome distribution shifts and be extensible for both source/target domains.
3) Due to the lack of generalized semantic expressions that can be applied to various PolSAR data, it is difficult to realize the zero-shot learning (ZSL) and generalized ZSL (GZSL) framework to recognize unseen classes in PolSAR interpretation [30].

For the PolSAR CDS problems, Wang *et al.* [29] analyzed characteristic differences of Gaofen-3, Radarsat-2, and ALOS 2-PALSAR2 data at same sites, and similar work has been presented in [5] to show the effects of the PolSAR CDS problem. With various labeled samples from the source domain and a small number of labeled samples from the target domain, Fang *et al.* [7] and Sun *et al.* [31] proposed semisupervised domain adaptation for cross-domain PolSAR classification. Even fewer uDA methods have been developed for SAR interpretation, Dong *et al.* [32] first proposed component-ratio-based distance (CRD) for cross-source PolSAR classification; it classifies the unlabeled target domain to the same land covers in the source domain. Moreover, Qin *et al.* [33] proposed relational-based transductive transfer learning for PolSAR time-series images. Additionally, Zhang *et al.* [34] proposed an unsupervised multilevel domain adaptation for multiband SAR image classification.

In general, there is a lack of methods to identify unseen categories in unlabeled target-domain PolSAR data. Although unsupervised methods can classify certain land covers in PolSAR images [35], [36], their applications are limited by the model complexity and the uncertainty of the clusters. Supervised classification methods usually rely on specific labeled samples. However, accurately labeled samples are particularly difficult to obtain in PolSAR [3], [37]; it is necessary to further make the utmost of limited labeled samples in more unlabeled target domains and achieve a more robust and practical unsupervised cross-domain PolSAR classification framework. Inspired by these demands, to further improve the reuse efficiency of PolSAR labeled samples and expand application scopes, a uGZSDA is proposed in this article. The uGZSDA can infer new classes and seen classes in the unlabeled target domain only by the limited seen samples from the source domain. This research is of great significance to sample reuse, the generalization of PolSAR features, and scattering semantic expression of land covers for cross-domain PolSAR analysis. Limited studies have been conducted on SAR domain adaptations [3], [5], [11], and there are even no publications concerned on the ZSDA/GZSDA in PolSAR interpretation, to the authors' knowledge. The uGZSDA is a novel uDA framework that can classify and infer the unlabeled target domain with unseen classes based on limited seen samples from the source domain. The main contributions and advantages of this work are as follows.

1) A novel uDA framework, uGZSDA, is proposed for cross-domain PolSAR interpretation. It classifies and infers more abundant land cover types in the unlabeled target domain, reusing a few types of labeled samples from the source domain. It overcomes the limitation that the existing uDA methods can only deal with the case of same seen classes of source and target domains. The uGZSDA further extends the practical range of cross-domain PolSAR sample reuse.
2) This article develops the generalized semantic expressions for PolSAR typical land cover types, which are the inference basis of unseen categories. The proposed semantic expressions help to infer the land cover categories of unlabeled PolSAR data from cross-domain sources and scenes.
3) The performance of the proposed method has been validated in 16 datasets from Radarsat-2 and Gaofen-3, including seven types of land covers, which demonstrate the robustness and effectiveness of the proposed uGZSDA.

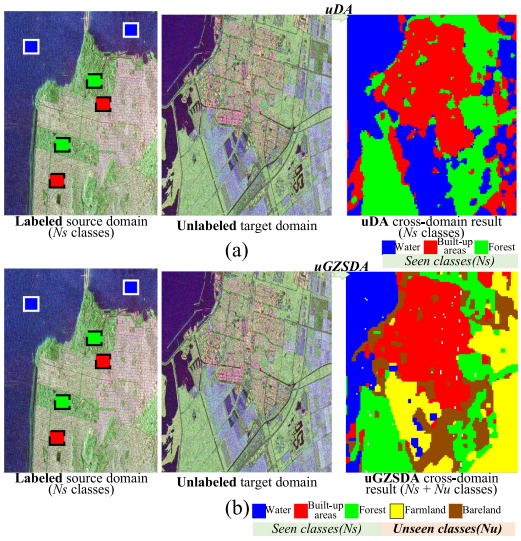


Fig. 2. Comparison of uDA and uGZSDA effects. (a) uDA case. (b) uGZSDA case.

The rest of this article is organized as follows. In Section II, the PolSAR image CDS problems and related works are described. In Section III, the proposed uGZSDA framework is presented. Section IV describes the datasets, the experimental setting, the experimental results, and discussion. Finally, Section V concludes this article.

II. uGZSDA Problems and Related Works

In this section, the uGZSDA problem formulation is first illustrated. Then, CDS of typical cross-domain PolSAR data is illustrated, followed by a brief overview of previous studies related to statistical scattering components [5].

A. uGZSDA Problem Formulation

Most of the existing uDA methods can only deal with the cross-domain case of the same categories between source and target domains [19], [22]. They cannot handle the newly acquired unlabeled target domain with more abundant categories than those in the labeled source domain [23], [38]. These prevent the reuse of labeled samples and the transfer of the training model. This article focuses on a more realistic cross-domain classification problem arising from real-world scenarios, where categories in the unlabeled target domain are greater than or equal to those in the source domain (seen classes). It is expected to recognize newly obtained unlabeled target-domain data, including any classes (seen and unseen classes), once the model is learned. This is a uGZSDA problem. The class-level semantic representations that are independent of specific images are usually indispensable in the uGZSDA for inferring unseen classes of the unlabeled target domain.

To facilitate our presentation in the following sections, Fig. 2 further illustrates the differences between the common uDA and the proposed uGZSDA. The uGZSDA deals with both the seen and unseen classes in cross-domain scenarios. These can not only reuse the labeled samples from the source domain, but also make the target-domain results more in line with the actual scenarios and needs. More specifically, the uGZSDA problem is formulated as follows. Given a labeled dataset $\mathbf{D}_S = (\mathbf{x}_{Si}, y_{Si})$, $i = 1 : n_S$, from the source domain $\mathbf{S}$, $y_{Si} \in y_S$, y_S denotes the seen classes' label space in the source domain. For the unlabeled target domain $\mathbf{D}_T = \mathbf{x}_{Tj}$, $j = 1 : n_T$, which contains both the seen and unseen classes. The goal of the uGZSDA is to predict target labels y_T (the class types are more than y_S, $y_S \in y_T$), based on the reuse of labeled source training data and semantic representations. The uGZSDA integrates the problems in both uDA and GZSL: the domain shift problems between $\mathbf{D}_S$ and $\mathbf{D}_T$, and how to infer the unseen classes in $\mathbf{D}_T$. Typical uDA cannot realize the unseen classes inferring, and GZSL usually cannot overcome the cross-domain shifts.

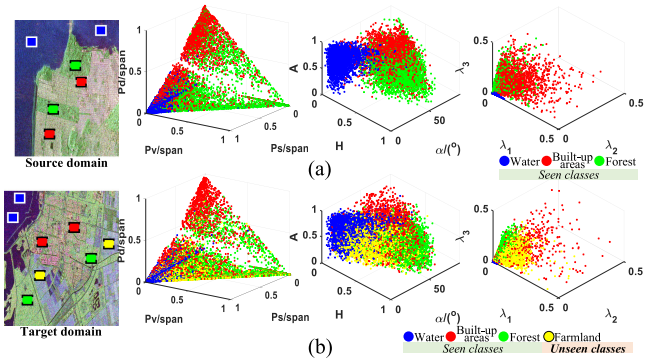


Fig. 3. Domain shifts between source and target domains in uGZSDA setting. (a) Source-domain data and corresponding typical polarimetric features. (b) Target-domain data and corresponding typical polarimetric features.

B. CDS Between Cross-Domain PolSAR Data

Fig. 3 shows the CDS problem in uGZSDA setting by three groups of typical polarimetric features, including Yamaguchi decompositions ($Ps/span$, $Pv/span$, and $Pd/span$), H/alpha/A decompositions (H, α, and A), and eigenvalues of coherency matrix (λ_1, λ_2, and λ_3). There are large distribution deviations of the corresponding categories in cross-domain PolSAR data and certain intersections between different categories. These CDS problems seriously hinder valuable labeled samples reuse among cross-domain data [5]. In addition, there are strong interferences and interleaving between the unseen class features in the target domain and seen class features in the source domain. That is, it is infeasible to identify the new classes of the target domain by samples from the source domain. To realize the uGZSDA, not only the CDS problem between the source and target domains should be overcome, but also the unseen classes should be inferred according to the polarimetric distribution characteristics and semantic spaces [24], [30].

C. Statistical Scattering Components

In our previous research [5], [11], statistical scattering components (SSCs) are proposed as the feature base for cross-domain PolSAR interpretation. Based on the widely applicable polarimetric scattering mechanism, the SSC obtains the histogram

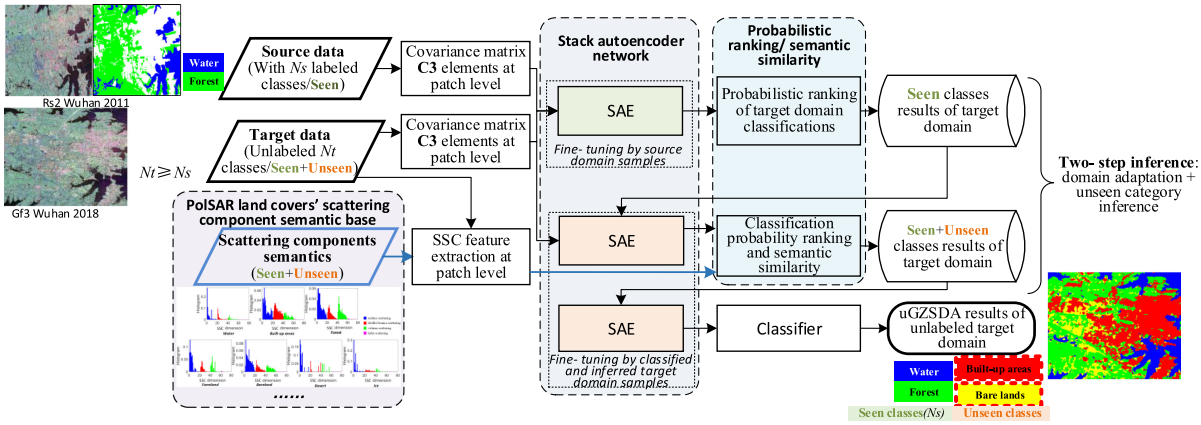


Fig. 4. Flowchart of the proposed uGZSDA framework.

statistics of land covers' scattering components (e.g., Yamaguchi decompositions) at the patch level. Thus, the SSC can effectively associate the scattering and statistical component information with the essential attributes of PolSAR land covers. It performs well for cross-domain PolSAR applications, including, but not limited to, Gaofen-3, Radarsat-2, AIRSAR, Pi-SAR, UAVSAR, and Sentinel-1. As a feature paradigm expression, the SSC is acquired in an unsupervised way, including polarimetric decompositions extraction, Wishart clustering, histogram statistics, and dimensionality reduction. The SSC helps to realize uDA in cross-domain PolSAR tasks.

Both theoretical and practical verifications show that the SSC expression is associated with the essential scattering and statistical characteristics of PolSAR land covers. This unsupervised feature expression manner is used to form the semantic expression of PolSAR land covers. The semantic expression of typical PolSAR land covers can be obtained from large-scale cross-domain databases based on the SSC manner. These scattering component semantics (SCSs) are also the semantic supports of uGZSDA inference.

III. Proposed Method

As shown in Fig. 4, the proposed uGZSDA framework of cross-domain PolSAR classification is shown.

In this framework, the main land cover classes of the source domain can be fewer than or the same as those of the target domain, but the target domain does not contain any labeling information. First, the SAE network is used to extract source/target-domain features, and the SCS expressions of typical land covers are constructed by the large-scale cross-domain databases and the SSC manner. Second, when combining the aligned network features and source-domain labels, the most likely seen class samples in the target domain are classified by probability sorting. The SAE network is retrained by using the obtained labeled seen samples from the target domain. Third, combined with the retrained SAE network and target-domain seen classes, the most likely unseen class samples in the target domain are inferred by using the classification probability and semantic similarity distances, and the specific unseen category is inferred by combining the semantic similarities with SCS.

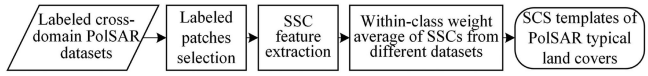


Fig. 5. Process of constructing SCS for typical PolSAR land covers.

Finally, the selected seen and inferred unseen class samples in the target domain are used to further retrain the SAE network; the retrained network features and the classifier are applied to further classify the target domain. The detailed descriptions of this unsupervised cross-domain classification and inference are as follows.

A. Scattering Component Semantics

At present, the semantic information used in remote sensing image ZSL and GZSL mainly comes from Word2Vec or SUN attributes [24], [30]. However, these semantics from natural languages and natural image models are not befitting for PolSAR description. Since SAR mechanisms are generally described by scattering characteristics, the effects of the Word2Vec or SUN attributes semantics for PolSAR land cover descriptions are limited or invalid [30], [39]. To obtain the more befitting semantic expression of PolSAR land covers, this section attempts to construct semantic expression of PolSAR land covers through large-scale cross-domain databases and the SSC manner. The obtained SCS is the inferring basis for the uGZSDA framework.

Combined with labeled cross-domain PolSAR databases and SSC (including Yamaguchi decompositions extraction, Wishart clustering, histogram statistics, and dimensionality reduction) [5], [11], the process of creating SCS for typical PolSAR land covers mainly includes four steps: labeled patch selection from cross-domain databases, SSC feature extraction of each patch, within-class weight average of SSCs, and the SCS templates are formed by the averaged SSCs. Among these steps, the labeled patch selection means to select the patches with high confidence (combined with the corresponding annotations, the homogeneous patches are selected). The within-class weight average step refers to the weighted average of SSCs of the same type of land covers in different datasets. Considering the SCS's universality for the dataset from different PolSAR sensors, the

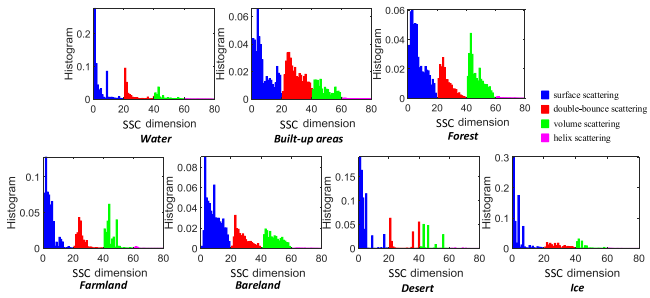


Fig. 6. SCS of typical PolSAR land covers.

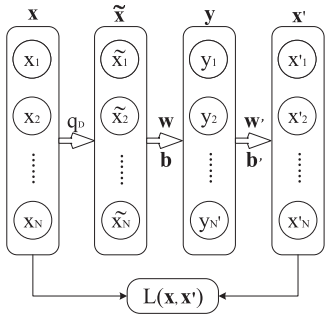


Fig. 7. Illustrations of the DAE basic block.

weights are set the same, and the sample sizes of every land cover from each dataset are fixed.

Furthermore, Fig. 6 illustrates the quantized SCS visualization effect of typical land covers obtained according to the flow in Fig. 5. In these semantic expressions, the similarities between water and ice and between forest and farmland are high, which are well consistent with the basic facts. Theoretically, not only the land covers in Fig. 6 can be described by SCS, but also the other typical land covers can be expressed in this way when there are enough corresponding labeled samples.

B. Feature Extraction Model Based on the SAE

Deep learning methods have been widely applied in SAR data interpretation [40]. As a typical neural network, the autoencoder (AE) network can automatically learn features from unlabeled data and give better feature descriptions than the original data [41]–[43]. The AE tries to find a representation close to the input, and some hidden features can be found in this process. The stacked autoencoder (SAE) model is a deep neural network model composed of multilayer AEs, and the output of the former layer is used as the input of the latter layer, and the output representation of the highest level can be used as input to a stand-alone supervised learning algorithm [42], [43]. It is widely used in unsupervised feature extraction and expression of remote sensing image classification [15], [44], [45]. SAE has shown great advantages and attraction in unsupervised PolSAR feature extractions in recent years [44], [46]. The commonly used SAE models in SAR interpretation include stacked sparse AE and stacked denoising autoencoder (DAE). As one of the typical and mature models, the stacked DAE can select effective feature representation from high-dimensional features in case of without labels; therefore, it can effectively reduce the cost of sample labeling [42], [43]. Combined with the above characteristics, the stacked DAE is applied.

As shown in Fig. 7, the DAE is the basic block of the applied SAE, which incorporates the entire n inputs. For a DAE, given a normalized N-dimensional input vector $\mathbf{x}$, an N-dimensional hidden layer $\mathbf{y}$ is generated as follows [42]:

$$\mathbf{y} = s(\mathbf{wx} + \mathbf{b}) = \mathbf{f}_\theta(\mathbf{x}) \tag{1}$$

where $\mathbf{w}$ is the weight matrix and $\mathbf{b}$ is the bias vector. θ is the parameter set $\{\mathbf{w}, \mathbf{b}\}$, and s is referred to as the sigmoid function. $\mathbf{y}$ is, then, mapped back as the DAE output $\mathbf{x}'$

$$\mathbf{x}' = s\left(\mathbf{w}'\mathbf{y} + \mathbf{b}'\right) = \mathbf{g}'_\theta(\mathbf{y}) \tag{2}$$

where $\mathbf{b}'$ is a different bias vector. The reverse weight matrix $\mathbf{w}'$ can optionally be constrained by $\mathbf{w}' = \mathbf{w}^T$, and thus, the DAE has tied weights. Through the minimization of a loss function $\mathbf{L}$ that represents the error between $\mathbf{x}$ and $\mathbf{x}'$ by using the stochastic gradient descent algorithm, θ and θ' can be determined [42]. The well-known squared error is applied in the loss functions

$$\mathbf{L}\left(\mathbf{x}, \mathbf{x}'\right) = ||\mathbf{x} - \mathbf{x}'||^2. \tag{3}$$

More specifically, the covariance matrixes $\mathbf{C3}$ of source- and target-domain partitioned samples at the patch level are the input $\mathbf{x}$ of the SAE. The partitioned samples patch size $n \times n$ and the overlapping rate of patches are optional. (Specifically, the optional range of n is 20–60. To improve efficiency, the default and empirical value of n can be set to 20, and the overlap rate is 50%.) For each partitioned sample, the amplitudes of $\mathbf{C3}$ elements are used to put into the SAE for feature extraction. Due to the symmetry of $\mathbf{C3}$ elements, the input dimension of each patch is $n \times n \times 6$. In the SAE step, the input vector $\mathbf{x}$ is first corrupted into a noisy version $\tilde{\mathbf{x}}$ by means of a stochastic mapping q_D, as shown in Fig. 7. Then, the DAE structure generates the output features $\mathbf{x}'$. The loss function $\mathbf{L}$ is established based on $\mathbf{x}$ and $\mathbf{x}'$, according to (3).

In the SAE structure, each DAE process above is regarded as a pretraining stage, and a fine-tuning stage is needed after it [42]. In the pretraining stage, the input vectors are trained one after another according to the SAE structure, and the final optimal parameters are determined in the fine-tuning process. Finally, the corresponding feature representations can be acquired by pretraining and fine-tuning stages. The number of SAE layers is adaptively selected according to the training sets from the source domain and the target domain.

The SAE has been used three times in the uGZSDA framework, and the specific process is illustrated in Fig. 8. In the first round of feature extraction, SAE1 is fine-tuned by the labeled samples from the source domain. The seen classes' samples in the target domain can be classified through the SAE features and probability sorting. In the second round, the obtained seen classes' samples from the target domain are used to retrain the SAE2 network. The unseen classes' samples in the target domain are classified through above retrained SAE2 features, probability sorting, and SCS semantic similarity. Third, the obtained seen

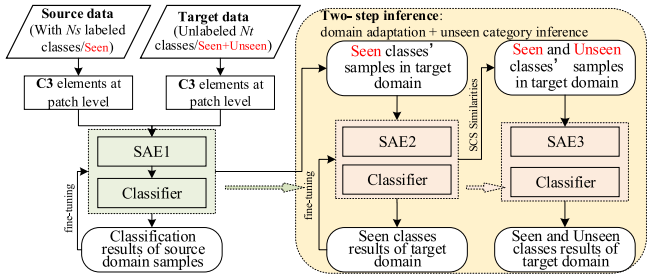


Fig. 8. SAE usage progress in the uGZSDA.

and unseen classes' samples are input to the SAE3 network for feature extraction, and the classifier is retrained to further classify the whole target domain.

C. Probabilistic Ranking and Semantic Similarity

For different source- and target-domain data, the optimized feature dimensions obtained by the SAE are also different, while the dimension of the constructed SCS is uniform. In addition, the types of seen classes for training are usually different from those in the target domain, and the differences between different SCSs are also large. If the mapping between image features and SCS is obtained by fixed mapping matrixes [49], under the condition of insufficient training classes and unbalanced seen and unseen classes, the projection domain shift problem [50] may easily occur and reduce the uDA effect. To match the variability of SAE feature dimensions, based on the SAE features and the created SCSs, a two-step inference is applied to classify and infer the unlabeled target domain by reusing limited labeled samples from the source domain. The two-step inference mainly depends on the classification probability sorting and semantic similarity, and the details are described in Algorithm 1.

1) Seen Class Classification in the Target Domain: Based on SAE1, the features of the unlabeled target domain are extracted, and the target domain is classified into seen land covers. Although SAE1 is trained by the samples from source and target domains, the fine-tuning is based on the labeled samples from the source domain. Therefore, it is inevitable that the results of obtained target-domain seen classes are still affected by the domain shifts. To reduce the influence of domain shifts, classification probability ranking is introduced.

For the initial classification results of the target domain, the probabilities of each seen class are sorted. Only the first p samples with larger probabilities are selected as corresponding determined seen samples (p is adjustable), denoted as $\mathbf{x}_T^{\text{Seen}}$. In addition, to retrain SAE2 to make it more suitable for the target domain, unlabeled samples $\mathbf{x}_T^{\text{unlabeled}}$ are randomly selected from the target domain. $\mathbf{x}_T^{\text{Seen}}$ and $\mathbf{x}_T^{\text{unlabeled}}$ are used to retrain SAE2, and the retrained SAE2 is further fine-tuned based on the obtained $\mathbf{x}_T^{\text{Seen}}$. SAE2 is used to extract features from the target domain and classify it, and the results with the larger probabilities are determined as the labeled seen samples of the target domain, i.e., update $\mathbf{x}_T^{\text{Seen}}$.

2) Unseen Class Inference in the Target Domain: Based on SAE2, only seen classes in the target domain can be classified.

Algorithm 1: Two-Step Inference Algorithm.

Input: The labeled source domain $\mathbf{D}_S$ with N_s seen classes (each class with p samples), the SCS templates, the unlabeled target domain $\mathbf{D}_T$, the unseen class number N_u ($N_s \geq 2$, $N_u \geq 0$, $N_t = N_s + N_u$).
Seen class classification in $\mathbf{D}_T$:
for each samples in $\mathbf{D}_S$ and $\mathbf{D}_T$ **do**
  Extract the amplitudes of **C3** features of train samples as input of SAE1
end for
Training the SAE1 by unlabeled samples from $\mathbf{D}_S$ and $\mathbf{D}_T$
Fine-tuning the SAE1 by labeled samples from $\mathbf{D}_S$
Class the $\mathbf{D}_T$ by SAE1 features, labeled samples of $\mathbf{D}_S$, and classifier.
The first p samples of each class with larger probabilities are selected as corresponding determined labeled seen sample set (p is adjustable) in $\mathbf{D}_T$, denote as $\mathbf{x}_T^{\text{Seen}}$
SAE2 is trained by $\mathbf{x}_T^{\text{Seen}}$ and unlabeled samples in $\mathbf{D}_T$
$\mathbf{D}_T$ is reclassified by SAE2 features, $\mathbf{x}_T^{\text{Seen}}$ and classifier, update $\mathbf{x}_T^{\text{Seen}}$ by the updated probabilistic ranking.
Unseen class inference in $\mathbf{D}_T$:
for each samples in $\mathbf{D}_T$ **do**
  SSC features are extracted
  The similarities between each semantic template SCS and samples' SSC features are calculated, select the largest similarity as the inferred class of the sample.
end for
The number of labels corresponding to each class is counted, removing seen classes in inferred results, and the remaining classes are sorted from large to small, the first N_u classes of the sorted numbers are determined as unseen classes.
for Each unseen class $k \leq N_u$ **do**
  The p unseen samples with high semantic similarity and low classification probability in step 11 are selected as corresponding determined unseen sample set in $\mathbf{D}_T$, denote as $\mathbf{x}_T^{\text{Unseen}}$.
end for
$\mathbf{x}_T^{\text{Seen}}$ and $\mathbf{x}_T^{\text{Unseen}}$ are used to train and fine-tune SAE3.
The $\mathbf{D}_T$ is re-classified by SAE3 features, $\mathbf{x}_T^{\text{Seen}}$, $\mathbf{x}_T^{\text{Unseen}}$ and classifier.
Output: Classified and inferred results (y_T) for the target domain $\mathbf{D}_T$ with N_t classes.

To further inference unseen classes, for each unlabeled sample in the target domain, their SSC [5] features are extracted, and the similarities between each semantic template SCS and samples' SSC features are calculated. The similarity vector $\mathbf{Sim}_{\text{Sem}}$ is expressed by cosine similarity [47], [48]

$$\mathbf{Sim}_{\text{Sem}} = [\cos(\mathbf{s}_T, \mathbf{S}_1), \ldots, \cos(\mathbf{s}_T, \mathbf{S}_i), \ldots, \cos(\mathbf{s}_T, \mathbf{S}_{N_c})] \tag{4}$$

TABLE I
DATA DESCRIPTIONS AND USAGES OF THE APPLIED POLSAR DATASET

Dataset name	Sensor	Imaging time	Size(pixels)	Imaging areas	Main land cover types	GT	Usage
Rs2-Sanf	Radarsat-2	Apr. 2008	1600 × 1200	San Francisco, USA	Water, Built-up areas, Forest	Yes	SCS[1] & uGZSDA[2]
Rs2-Vanc	Radarsat-2	Apr. 2008	3000 × 1000	Vancouver, CAN	Water, Built-up areas, Forest	Yes	SCS
Rs2-Vanco	Radarsat-2	Apr. 2008	2600 × 1600	Vancouver, CAN	Water, Built-up areas, Forest, Farmland	Yes	SCS,uGZSDA
Rs2-Flev	Radarsat-2	Apr. 2008	1400 × 1200	Flevoland, NED	Water, Built-up areas, Forest, Farmland	Yes	SCS,uGZSDA
Rs2-Flevo	Radarsat-2	Apr. 2008	1400 × 1200	Flevoland, NED	Water, Built-up areas, Forest, Farmland, Bareland	Yes	uGZSDA
Rs2-Wuh	Radarsat-2	Dec. 2011	3000 × 2800	Wuhan, CHN	Water, Built-up areas, Forest, Bareland	Yes	SCS,uGZSDA
Gf3-Sanf	Gaofen-3	Sep. 2017	1150 × 1800	San Francisco, USA	Water, Built-up areas, Forest	Yes	SCS
Gf3-Wuh	Gaofen-3	Aug.2017	3500 × 2000	Wuhan, CHN	Water, Built-up areas, Forest, Farmland	Yes	SCS
Gf3-Wuhan	Gaofen-3	Dec. 2018	3800 × 4500	Wuhan, CHN	Water, Built-up areas, Forest, Bareland	Part	SCS,uGZSDA
Gf3-JYH	Gaofen-3	May. 2017	650 × 1190	Wuhan, CHN	Water, Built-up areas, Forest	Yes	SCS
Gf3-Yingk	Gaofen-3	Jan. 2018	1400 × 2000	Yingkou, CHN	Water, Built-up areas, Ice	Part	SCS,uGZSDA
Gf3-Yingko	Gaofen-3	Jan. 2018	1100 × 2500	Yingkou, CHN	Water, Built-up areas, Ice	Part	SCS
Gf3-Dunh	Gaofen-3	Jun. 2018	750 × 900	Dunhuang, CHN	Built-up areas, Desert	Yes	SCS
Gf3-Jiayg	Gaofen-3	Jul. 2017	1200 × 1800	Jiayuguan, CHN	Built-up areas, Desert, Farmland	Part	SCS,uGZSDA
Gf3-Sheny	Gaofen-3	Oct. 2017	1600 × 2000	Shenyang, CHN	Water, Built-up areas, Forest, Farmland	Part	SCS
Gf3-WuhA	Gaofen-3	Apr.2017	600 × 970	Wuhan, CHN	Water, Built-up areas, Forest, Farmland	Part	uGZSDA

[0]SCS[1] represents the usage of creating SCS and verifying SCS effectiveness. uGZSDA[2] indicates the usage of uGZSDA framework verification.

where $\mathbf{s}_T$ represents the SSC features of each unlabeled samples; $\mathbf{S}_1, \mathbf{S}_i, \ldots, \mathbf{S}_{N_c}$ are the SCS templates corresponding to the seen and unseen classes (N_c is the number of the SCS templates; it is usually larger than the concerned N_t in the target domain).

For each instance in unlabeled $\mathbf{D}_T$, the largest similarity in $\mathbf{sim}_{\text{Sem}}$ is selected as the inferred class of this instance; the inferred result is denoted as N_i $(2 \leq i \leq N_c)$. For all the inferred results of instances in $\mathbf{D}_T$, the number of labels corresponding to each class is counted, denoted as $\mathbf{Sum}_N$. The classes in $\mathbf{Sum}_N$ are consistent with the number of classes in the semantic base. Based on $\mathbf{x}_T^{\text{Seen}}$, the seen classes in $\mathbf{Sum}_N$ are removed. The remaining classes are sorted from large to small; the first N_u classes of the sorted numbers are determined as unseen classes, which are denoted as y_U. For each class in y_U, the first p samples with high semantic similarity and low classification probability in step (11) are selected to form the labeled sample set $\mathbf{x}_T^{\text{Unseen}}$ for unseen classes in $\mathbf{D}_T$

$$\mathbf{Sum}_N = [\text{Sum}_{N_1}, \text{Sum}_{N_2}, \ldots, \text{Sum}_{N_i}, \ldots, \text{Sum}_{N_c}]. \quad (5)$$

Based on the updated $\mathbf{x}_T^{\text{Seen}}$ and the $\mathbf{x}_T^{\text{Unseen}}$ obtained by semantic inference, SAE2 is retrained again to form SAE3. Because SAE3 is trained by target-domain samples and the fine-tuning process also depends on the target-domain samples, SAE3 is suitable for the target domain to obtain the seen and unseen classes at the same time. After three rounds of training and fine-tuning, the SAE is changed from depending on the source domain to the target domain; the obtained features are more suitable for the target domain. Thus, the domain-shift problem is reduced in the SAE training and fine-tuning.

IV. EXPERIMENTS AND RESULT ANALYSIS

The applied dataset and cross-domain settings are first introduced in this section. Then, the proposed SCS and the uGZSDA are verified. Moreover, further comparisons are presented.

A. Dataset Descriptions and Experimental Settings

The effectiveness of the proposed method was tested on six Radarsat-2 spaceborne PolSAR data and ten Gaofen-3 spaceborne PolSAR data. The basic information of selected experimental data is shown in Table I. These cross-domain data include the cross-sources, cross-scenes, cross-temporal data, and the types of land covers are different. Seven typical types of land covers are contained in these datasets. For each land cover type in dataset for creating SCS, 500 samples are randomly selected to construct corresponding SCS templates. According to the setting of the uGZSDA in Section II-A, the detailed types of unseen classes are unknown, but only the number of unseen classes N_u can be preset.

The cross-domain units in Table II are set to verify uGZSDA effectiveness. To illustrate the land cover differences between the source and target domains, the corresponding seen and unseen land covers are shown in advance. Actually, the unseen classes in the target domain are unknown, they are inferred by combining with SCS. If a different number of unseen classes are given, different classification and inference results can be obtained. By applying the uGZSDA, labeled samples from a source domain and generalized SSC can be reused in many other target-domain data with different land covers. The training samples from the source domain are selected randomly. For each cross-domain unit, the number of training samples of each seen class is set to 1000 by defaults, and the number of training samples should be less than 10% of the number of source-domain samples.

B. SCS Performances

The SCS is constructed based on labeled cross-domain datasets. These generalized semantic templates can support different unlabeled PolSAR data for land cover inference. To verify its generalization and effectiveness, three aspects of experiments are conducted. First, for the obtained SCS templates, the correlations between each kind of SCS are obtained according to cosine similarity, as shown in Fig. 9. The correlations between SCSs

TABLE II
BASIC EXPERIMENTAL UNITS INTRODUCTION FOR UGZSDA CROSS-DOMAIN CLASSIFICATION AND INFERENCE

Cross-domain Units	**Source domain and Seen classes**		**Unlabeled Target domain**		
Source-target	Source domain	Seen classes	Target domain	Seen classes	**Unseen classes(by inference)**
Rs2-Sanf-c3/Rs2-Flev-c4	Rs2-Sanf	Water, Built-up areas, Forest	Rs2-Flev	Water, Built-up areas, Forest	**Farmland**
Rs2-Sanf-c3/RFlevo-c5	Rs2-Sanf	Water, Built-up areas, Forest	Rs2-Flevo	Water, Built-up areas, Forest	**Farmland,Bareland**
Rs2-Sanf-c2/RFlev-c4	Rs2-Sanf	Water, Forest	Rs2-Flev	Water, Forest	**Built-up areas,Farmland**
Rs2-Sanf-c2/Rs2-Flevo-c5	Rs2-Sanf	Water, Forest	Rs2-Flevo	Water, Forest	**Built-up areas,Farmland,Bareland**
Rs2-Sanf-c3/Rs2-Vanco-c4	Rs2-Sanf	Water, Built-up areas, Forest	Rs2-Vanco	Water, Built-up areas, Forest	**Farmland**
Rs2-Sanf-c3/Gf3-Wuhan-c4	Rs2-Sanf	Water, Built-up areas, Forest	Gf3-Wuhan	Water, Built-up areas, Forest	**Bareland**
Rs2-Sanf-c2/Gf3-Yingk-c4	Rs2-Sanf	Water, Built-up areas	Gf3-Yingk	Water, Built-up areas	**Bareland,Ice**
Gf3-Dunh-c2/Gf3-Jia-c3	Gf3-Jiayg	Desert, Built-up areas	Gf3-Jiayg	Desert, Built-up areas	**Farmland**
Rs2-Sanf-c2/Gf3-WuhA-c4	Rs2-Sanf	Water, Built-up areas	Gf3-WuhA	Water, Built-up areas	**Forest,Farmland**

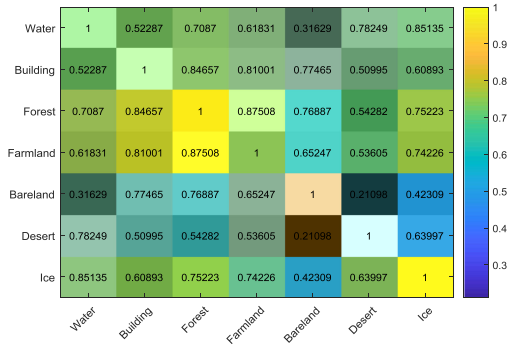


Fig. 9. Correlations between typical land covers' SCS templates.

are in line with the characteristics of PolSAR land covers. For example, the SCS relations between water and ice and between water and desert have high correlativity, because they are mainly surface scattering targets. The similarity between forest SCS and farmland SCS is high because they are both vegetation land covers, including surface scattering and volume scattering. The similarity between built-up areas SCS and forest SCS is also high; both of them contain dihedral scattering. In addition, in Fig. 9, there are great semantic differences between SCSs of built-up areas and water, built-up areas and desert, desert and forest, desert and farmland, and farmland and bareland, which are also in line with PolSAR land cover's scattering characteristics. These indicate that the SCS templates are suitable for the expression of PolSAR typical land covers.

Second, the SSC features of each data sample are directly matched with each SCS template, and the nearest template is taken as the inferred result of the sample. The SCS template matching results of the four datasets in Table I are shown in Fig. 10. From the results and corresponding ground truth (GT), it can be seen that the SCS templates can be applied to different PolSAR data and can also help to infer the main land cover information. However, due to the CDS problem, there are some wrong inference results when only using the SCS templates. For example, the difference between ice and desert SSC is not significant enough, which leads to some wrong inference results in the third row of Fig. 10.

Third, the SCS template matching result of the Rs2-Flev dataset is shown in Fig. 11, and as a comparison, the results

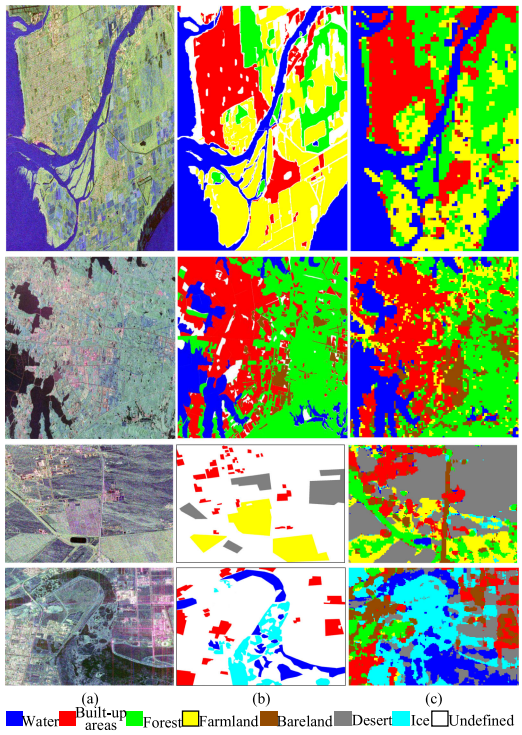


Fig. 10. Examples of the SCS template matching results, (a) Pauli images of Rs2-Vanco, Rs2-Wuh, Gf3-Jiayg, and Gf3-Yingk datasets. (b) GT of corresponding datasets. (c) SCS matching results.

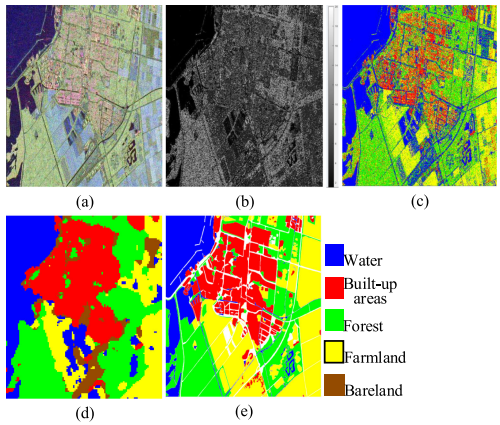


Fig. 11. SCS template matching result of the Rs2-Flev dataset and comparisons. (a) Pauli image. (b) unsupervised Yamaguchi–Wishart clustering results. (c) Artificially assigned results of (b). (d) SCS matching results. (e) GT.

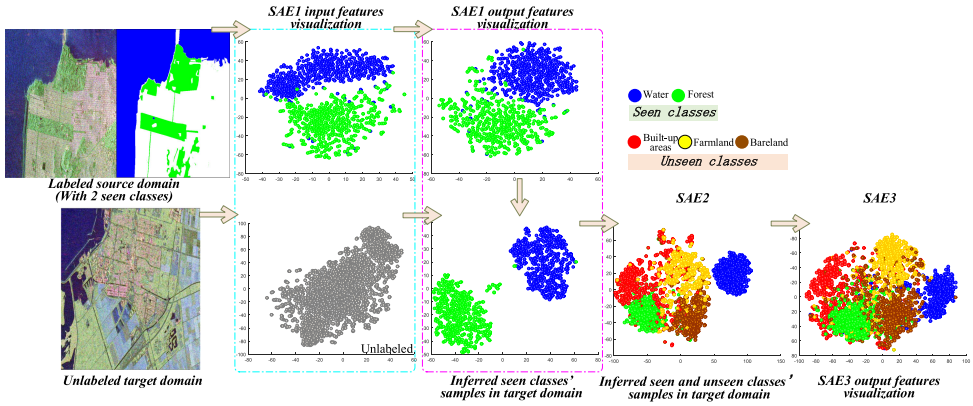


Fig. 12. t-SNE visualizations of the Rs2-Sanf-c2/Rs2-Flevo-c5 uGZSDA process.

obtained by unsupervised Yamaguchi–Wishart [51] are presented [the clusters of unsupervised results should be artificially assigned to land covers for evaluation in Fig. 11(c)]. Combined with GT, this group of experimental results can further show that the SCS templates are more flexible and factual. On one hand, SCS can help to obtain detailed land covers types for the unlabeled target-domain data. On the other hand, the SCS templates are generalized; they can be applied for different cross-domain PolSAR data.

C. uGZSDA Process Visualization

In general, the above semantic correlations and inference results are in line with expectations. The inference results in Fig. 10 are more abundant than GT; this is also consistent with the actual situation. For some complex scenes in Fig. 10, it is very difficult to obtain accurate and fine-grained labels, and the results obtained by the SCS matching can accurately classify the land cover types in unlabeled data.

To further demonstrate the effectiveness of the uGZSDA inference, taking Rs2-Sanf-c2/Rs2-Flevo-c5 unit as an example, the inference process is visualized in Fig. 12. Among the $\mathbf{C3}$ features input into SAE1, only the source-domain samples contain labels. After SAE1 and SAE2, the seen samples in the target domain are classified and confirmed. Combined with probability ranking and semantic similarity, the unseen classes in the target domain are inferred. The inferred samples are used to classify the target domain. In addition, Fig. 13 visualizes the $\mathbf{C3}$ features and SAE network features. The original features have large intraclass differences, while after the uGZSDA, the intraclass aggregation is optimized. Figs. 12 and 13 prove that the uGZSDA can infer the unlabeled target domain.

D. Evaluation of Proposed uGZSDA Frameworks

According to the definition and settings of the uGZSDA, the experimental units in Table II are evaluated quantitatively. The number of seen classes in the source domain is two to three; the training samples from the source domain are selected randomly. For each class, the number of training samples is 1000. The target domain does not contain any labels, and the unseen classes are unknown. To make the SAE network suitable for the target domain, unlabeled target-domain samples are randomly selected to participate in the training, and the total number of samples is the same as that of the source domain.

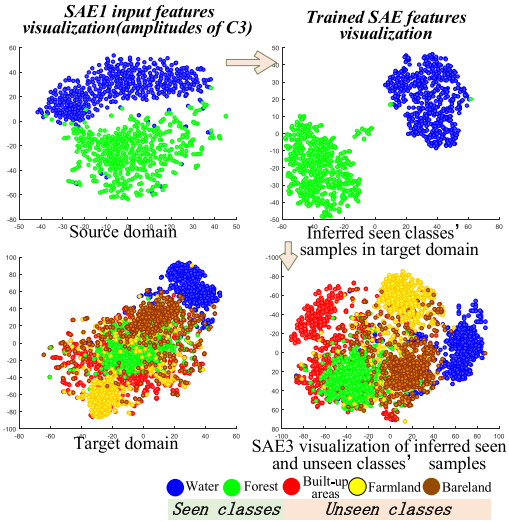


Fig. 13. t-SNE visualizations of the $\mathbf{C3}$ features and the SAE features.

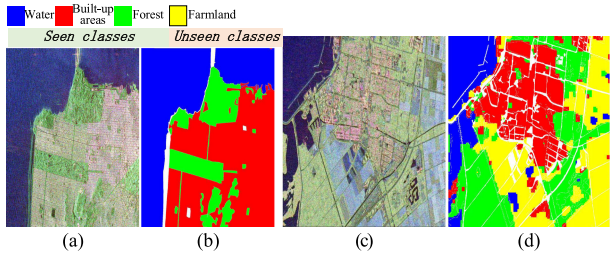


Fig. 14. uGZSDA results of Rs2-Sanf-c3/Rs2-Flev-c4 unit. (a) and (b) are the source-domain Pauli image and GT, respectively. (c) and (d) are the Pauli image and uGZSDA results of the target domain, respectively.

The uGZSDA results of eight units in Table II are shown in Figs. 14– 22. The target domains have no labeled samples in all experiments, and the categories in the target domain are different from the source-domain categories (unseen categories are unknown, but the number of unseen categories can be preset). From the results, by using two to three types of seen labeled

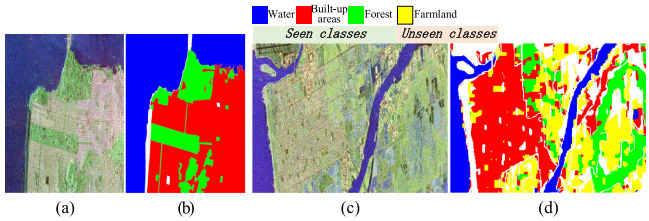


Fig. 15. uGZSDA results of Rs2-Sanf-c3/Rs2-Vanco-c4 unit. (a) and (b) are the source-domain Pauli image and GT, respectively. (c) and (d) are the Pauli image and uGZSDA results of the target domain, respectively.

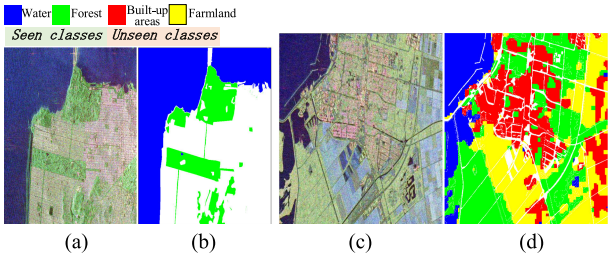


Fig. 16. uGZSDA results of Rs2-Sanf-c2/RFlev-c4 unit. (a) and (b) are the source-domain Pauli image and GT, respectively. (c) and (d) are the Pauli image and uGZSDA results of the target domain, respectively.

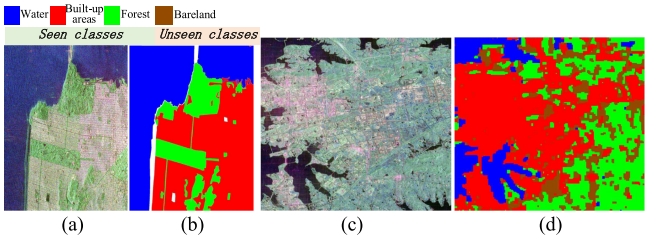


Fig. 17. uGZSDA results of Rs2-Sanf-c3/Gf3-Wuhan-c4 unit. (a) and (b) are the source-domain Pauli image and GT, respectively. (c) and (d) are the Pauli image and uGZSDA results of the target domain, respectively.

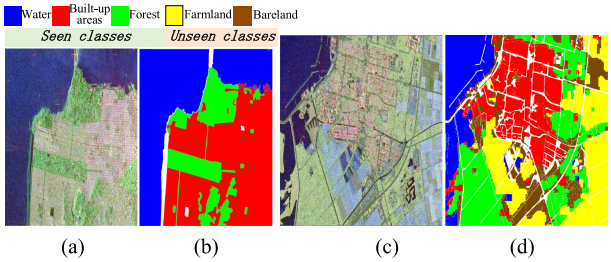


Fig. 18. uGZSDA results of Rs2-Sanf-c3/RFlevo-c5 unit. (a) and (b) are the source-domain Pauli image and GT, respectively. (c) and (d) are the Pauli image and uGZSDA results of the target domain, respectively.

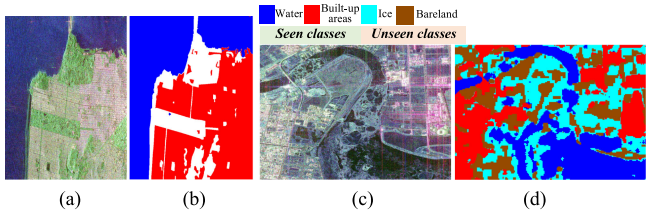


Fig. 19. uGZSDA results of Rs2-Sanf-c2/Gf3-Yingk-c4 unit. (a) and (b) are the source-domain Pauli image and GT, respectively. (c) and (d) are the Pauli image and uGZSDA results of the target domain, respectively.

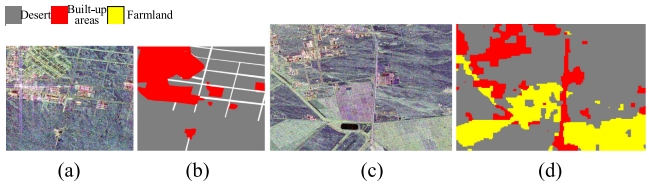


Fig. 20. uGZSDA results of Gf3-Dunh-c2/Gf3-Jia-c3 unit. (a) and (b) are the source-domain Pauli image and GT, respectively. (c) and (d) are the Pauli image and uGZSDA results of the target domain, respectively.

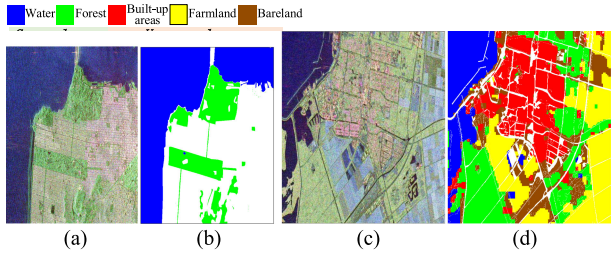


Fig. 21. uGZSDA results of Rs2-Sanf-c2/Rs2-Flevo-c5 unit. (a) and (b) are the source-domain Pauli image and GT, respectively. (c) and (d) are the Pauli image and uGZSDA results of the target domain, respectively.

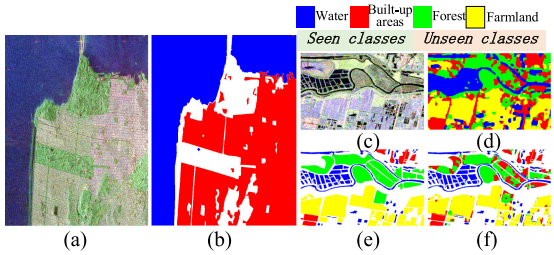


Fig. 22. uGZSDA results of Rs2-Sanf-c2/Gf3-WuhA-c4 unit. (a) and (b) are the source-domain Pauli image and GT, respectively. (c) and (d) are the Pauli image and uGZSDA results of the target domain, respectively. (e) and (f) are the partial GT and the corresponding results for evaluation, respectively.

TABLE III
EVALUATION OF RS2-SANF-C3/RS2-FLEV-C4 UGZSDA CROSS-DOMAIN RESULTS (%)

	Seen classes			**Unseen classes**
	Water	Built-up areas	Forest	**Farmland**
Water	83.79	5.44	5.95	4.81
Built-up areas	0.23	89.56	5.44	4.77
Forest	2.35	7.46	82.35	7.84
Farmland (Unseen)	3.57	4.87	10.23	**81.33**
				OA: **83.96**, Kappa: **0.7829**

samples from the source domain, combined with SCSs, four to seven types of land covers can be classified or inferred for the unlabeled target domain. The classification and inference results are mostly in line with the actual situations.

In particular, the quantitative evaluations of Rs2-Sanf-c3/Rs2-Flev-c4, Rs2-Sanf-c3/Rs2-Vanco-c4, Rs2-Sanf-c3/RFlevo-c5, Rs2-Sanf-c2/RFlev-c4, Rs2-Sanf-c2/Rs2-Flevo-c5, and Rs2-Sanf-c2/Gf3-WuhA-c4 units are shown in Tables III–VIII;

TABLE IV
EVALUATION OF RS2-SANF-C3/RS2-VANCO-C4 UGZSDA CROSS-DOMAIN RESULTS (%)

	Seen classes			Unseen classes
	Water	Built-up areas	Forest	**Farmland**
Water	92.67	2.20	0.00	5.13
Built-up areas	0.43	90.27	0.28	9.00
Forest	0.07	16.31	63.76	19.86
Farmland (Unseen)	1.53	10.87	11.11	**76.49**
OA: **81.36**, Kappa: **0.7424**				

TABLE V
EVALUATION OF RS2-SANF-C3/RFLEVO-C5 UGZSDA CROSS-DOMAIN RESULTS (%)

	Seen classes			Unseen classes	
	Water	Built-up areas	Forest	**Farmland**	**Bareland**
Water	83.96	4.00	2.18	0.54	9.06
Built-up areas	0.17	85.40	1.00	1.71	11.08
Forest	2.18	9.68	69.56	5.76	12.38
Farmland (Unseen)	3.42	2.28	4.70	**79.80**	9.76
Bareland (Unseen)	0.02	6.33	5.37	8.78	**79.28**
OA: **78.92**, Kappa: **0.7365**					

TABLE VI
EVALUATION OF RS2-SANF-C2/RFLEV-C4 UGZSDA CROSS-DOMAIN RESULTS (%)

	Seen classes		Unseen classes	
	Water	Forest	**Built-up areas**	**Farmland**
Water	85.96	5.67	4.16	4.21
Forest	2.68	83.66	3.62	10.03
Built-up areas (Unseen)	0.49	23.64	**72.88**	2.99
Farmland (Unseen)	0.40	9.34	21.73	**68.53**
OA: **76.85**, Kappa: **0.6870**				

TABLE VII
EVALUATION OF RS2-SANF-C2/RS2-FLEVO-C5 UGZSDA CROSS-DOMAIN RESULTS (%)

	Seen classes		Unseen classes		
	Water	Forest	**Built-up areas**	**Farmland**	**Bareland**
Water	84.53	4.61	3.50	0.51	6.59
Forest	2.69	80.72	4.08	3.04	9.03
Built-up areas (Unseen)	0.31	12.95	**79.00**	0.75	6.34
Farmland (Unseen)	0.50	14.83	9.50	**70.55**	4.58
Bareland (Unseen)	6.55	10.85	13.85	8.27	**60.26**
OA: **76.85**, Kappa: **0.7060**					

TABLE VIII
EVALUATION OF RS2-SANF-C2/GF3-WUHA-C4 UGZSDA CROSS-DOMAIN RESULTS (%)

	Seen classes		Unseen classes	
	Water	Built-up areas	**Forest**	**Farmland**
Water	95.81	1.11	3.02	0.06
Built-up areas	0.46	95.08	2.74	1.72
Forest (Unseen)	2.23	32.40	**64.03**	1.35
Farmland (Unseen)	2.01	4.73	1.63	**91.63**
OA: **83.51**, Kappa: **0.7696**				

the accuracies of the three cross-domain units can reach 76–83.96%. The units of Rs2-Sanf-c3/RFlevo-c5 and Rs2-Sanf-c2/RFlevo-c5 use the two to three types of labeled samples from the source domain to infer the five types of land covers' results in the unlabeled target domain. Since the GTs of Gf3-Wuhan, Gf3-Yingk, and Gf3-Yingko datasets are incomplete, the uGZSDA classifies or infers all samples in the target domain; therefore, these inference results in Rs2-Sanf-c2/Gf3-WuhA-c4, Rs2-Sanf-c2/Gf3-Yingk-c4, and Gf3-Dunh-c2/Gf3-Jia-c3 units are not evaluated quantitatively. Above inferred results and quantitative evaluation show the effectiveness and extensibility of the proposed uGZSDA method. Combined with semantic information, a small number of labeled samples in the source domain can play a better role in many unlabeled target-domain data with different land covers.

In addition, Figs. 17, 19, and 20 show the results of Rs2-Sanf-c3/Gf3-Wuhan-c4, Rs2-Sanf-c2/Gf3-Yingk-c4, and Gf3-Dunh-c2/Gf3-Jia-c3 cross-domain units, respectively. Because there is no complete GTs for these target domains, the quantitative evaluations of these three results are not available. The land cover scenarios of these three target-domain datasets are very complex, but the inference results are mostly consistent with the facts. For example, in Fig. 19, according to the built-up areas and water samples from the source domain, the built-up areas, water, ice, and bareland in the unlabeled target domain are obtained by inference, which are consistent with the characteristics of the corresponding region in imaging time. Compared with the results by SCS template matching in third and fourth rows of Fig. 10, the results in Figs. 19 and 20 are more accurate and applicable.

Especially, for the cross-domain unit Rs2-Sanf-c2/Gf3-WuhA-c4, the target-domain dataset Gf3-WuhA does not participate in SCS semantic creation, which can be considered as the newly acquired target-domain data without labels and prior information. Table VIII evaluates the accuracy of the results of cross-domain unit Rs2-Sanf-c2/Gf3-WuhA-c4 on a limited number of labeled samples. The results in Fig. 22 and Table VIII further prove the effectiveness of the proposed method. The above experiments show that the uGZSDA can fully utilize labeled samples in the source domain and semantic information, to overcome the CDS problem and classify the unlabeled target domain.

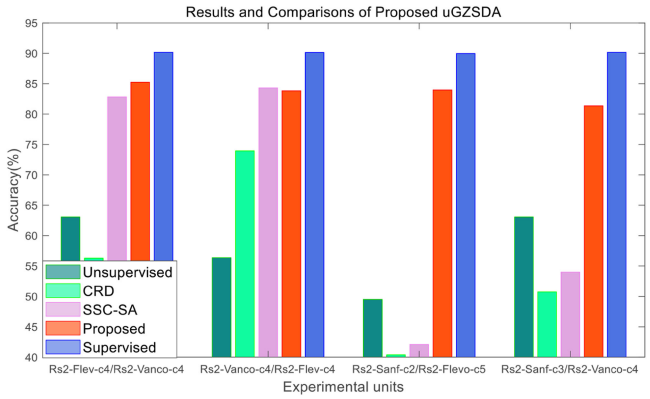


Fig. 23. Unsupervised cross-domain results based on the uGZSDA and comparisons.

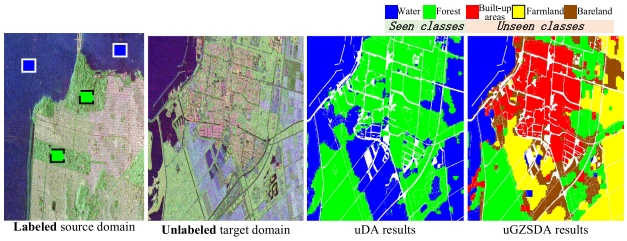


Fig. 24. Comparison of uDA and uGZSDA effects in the Rs2-Sanf-c2/Rs2-Flevo-c5 unit.

E. Comparisons and Special Cases

Since there are few well-established methods for the cross-domain PolSAR inference classification for the unlabeled target domain, it is difficult to get comparisons under the same experimental settings (the unlabeled target domain with more categories than that of the labeled source domain). However, because the number of unseen classes N_u can be set, the proposed uGZSDA model can degenerate to the conventional uDA when $N_u = 0$. To further compare with other typical uDA methods, N_u is set to 0, and the unsupervised cross-domain experiments are carried out using Rs2-Flev and Rs2-Vanco datasets. The labeled samples from the source domain include four classes. The Rs2-Vanco-c4/Rs2-Flev-c4 and Rs2-Flev-c4/Rs2-Vanco-c4 cross-domain results and comparisons are shown in Fig. 23. The applied comparisons include unsupervised Yamaguchi–Wishart clustering, supervised random forest methods in the target domain, CRD cross-domain method [32], and SSC-subspace alignment (SSC-SA) [5]. It is difficult to directly obtain the number and specific types of land cover by unsupervised clustering; therefore, we artificially assign the clusters to land covers for evaluation. In Fig. 23, the results of the uGZSDA($N_u = 0$) are comparable with those of SSC-SA, which are better than those of CRD and unsupervised results in the target domain.

In addition, to further highlight the differences between the uDA and the uGZSDA, the Rs2-Sanf-c2/Rs2-Flevo-c5 and Rs2-Sanf-c3/Rs2-Vanco-c4 cross-domain units are conducted and compared with the uDA, as shown in Figs. 23–25. The experiments show that the uGZSDA is more applicable than the uDA. The uGZSDA can not only make the labeled samples be reused

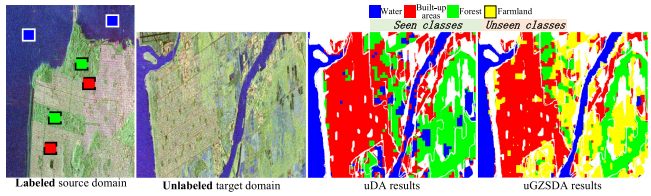


Fig. 25. Comparison of uDA and uGZSDA effects in the Rs2-Sanf-c3/Rs2-Vanco-c4 unit.

in cross-domain data, but also infer some new types of land covers in the unlabeled target domain, which further reduce the burden of labeling samples.

V. Conclusion

To further reuse the limited valuable labeled samples among cross-domain PolSAR images, the uGZSDA framework was proposed to classify and infer the seen and unseen classes in the unlabeled target domain based on generalized SCS semantics. In the uGZSDA framework, the domain shift problem between source and target domains can be reduced by three rounds of SAE fine-tuning. The seen and unseen classes in the unlabeled target domain were classified and inferred by using classification probabilities and semantic similarity distances with SCS templates. The uGZSDA realized cross-domain labeled samples reuse and overcame the limitation that the existing-domain adaptations can only deal with the case of same seen classes of source and the target domain. It further extended the practical range of cross-domain PolSAR sample reuse. Moreover, this article proposed and constructed the generalized semantic expression SCS for PolSAR typical land covers. The SCS helped to infer the unseen land covers of unlabeled PolSAR data from cross-domain sources or scenes. The proposed method was verified in seven types of land covers in 16 Radarsat-2 and Gaofen-3 datasets. By using two to three types of seen samples from the source domain, combined with generalized SCSs, the accuracies of four to five types of seen and unseen land covers in the unlabeled target domain can reach 76–83.96%.

The uGZSDA mainly focused on the labeled samples reuse from one source domain to multiple unlabeled target domains. In practice, labeled samples from multiple source domains may be available, and the categories in different source domains can be complementary for the unlabeled target domain with complex land covers. The follow-up research will focus on how to complement or weight the accurate labeled samples from multiple source domains, so as to provide more comprehensive labeled cross-domain samples for unlabeled target domains. This is also a way to further improve the cross-domain performance and PolSAR sample reuse application scope.

Acknowledgment

The authors would like to thank the Editor, Associate Editor, and anonymous reviewers for their constructive and helpful comments that greatly improve this article.

REFERENCES

[1] J.-S. Lee and E. Pottier, *Polarimetric Radar Imaging: From Basics to Applications*. Boca Raton, FL, USA: CRC Press, 2009.

[2] S. W. Chen and C. S. Tao, "PolSAR image classification using polarimetric-feature-driven deep convolutional neural network," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 4, pp. 627–631, Apr. 2018.

[3] Z. Huang, Z. Pan, and B. Lei, "What, where, and how to transfer in SAR target recognition based on deep CNNs," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 4, pp. 2324–2326, Apr. 2020.

[4] M. Rostami, S. Kolouri, E. Eaton, and K. Kim, "SAR image classification using few-shot cross-domain transfer learning, IEEE/CVF conference on computer vision and pattern recognition workshops," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops*, 2019, pp. 907–915.

[5] R. Gui, X. Xu, R. Yang, L. Wang, and F. Pu, "Statistical scattering component-based subspace alignment for unsupervised cross-domain PolSAR image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 7, pp. 5449–5463, Jul. 2021.

[6] B. Ren, Y. Zhao, B. Hou, J. Chanussot, and L. Jiao, "A mutual information-based self-supervised learning model for PolSAR land cover classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 11, pp. 9224–9237, Nov. 2021.

[7] Z. Fang, G. Zhang, Q. Dai, Y. Kong, and P. Wang, "Semisupervised deep convolutional neural networks using pseudo labels for PolSAR image classification," *IEEE Geosci. Remote Sens. Lett.*, to be published, doi: 10.1109/LGRS.2020.3036387.

[8] D. Tuia, C. Persello, and L. Bruzzone, "Domain adaptation for the classification of remote sensing data: An overview of recent advances," *IEEE Geosci. Remote Sens. Mag.*, vol. 4, no. 2, pp. 41–57, Jun. 2016.

[9] J. Geng, X. Deng, X. Ma, and W. Jiang, "Transfer learning for SAR image classification via deep joint distribution adaptation networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 8, pp. 5377–5392, Aug. 2020

[10] F. Ye, W. Luo, M. Dong, H. He, and W. Min, "SAR image retrieval based on unsupervised domain adaptation and clustering," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 9, pp. 1482–1486, Sep. 2019.

[11] R. Gui *et al.*, "A general feature paradigm for unsupervised cross-domain PolSAR image classification," *IEEE Geosci. Remote Sens. Lett.*, to be published, doi: 10.1109/LGRS.2021.3073738.

[12] W. Liu and R. Qin, "A multikernel domain adaptation method for unsupervised transfer learning on cross-source and cross-region remote sensing data classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 6, pp. 4279–4289, Jun. 2020.

[13] H.-W Jo *et al.*, "Deep learning applications on multitemporal SAR (Sentinel-1) image classification using confined labeled data: The case of detecting rice paddy in South Korea," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 11, pp. 7589–7601, Nov. 2020.

[14] X. Lu, T. Gong, and X. Zheng, "Multisource compensation network for remote sensing cross-domain scene classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 4, pp. 2505–2515, Apr. 2020.

[15] A. Elshamli, G. W. Taylor, A. Berg, and S. Areibi, "Domain adaptation using representation learning for the classification of remote sensing images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 9, pp. 4198–4209, Sep. 2017.

[16] Y. Li, Y. Zhang, X. Huang, and J. Ma, "Learning source-invariant deep hashing convolutional neural networks for cross-source remote sensing image retrieval," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 11, pp. 6521–6536, Nov. 2018.

[17] J. Guo, J. Yang, H. Yue, and K. Li, "Unsupervised domain adaptation for cloud detection based on grouped features alignment and entropy minimization," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5603413.

[18] O. Tasar, A. Giros, Y. Tarabalka, P. Alliez, and S. Clerc, "DAugNet: Unsupervised, multisource, multitarget, and life-long domain adaptation for semantic segmentation of satellite images," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 2, pp. 1067–1081, Feb. 2021.

[19] R. Adayel, Y. Bazi, H. Alhichri, and N. Alajlan, "Deep open-set domain adaptation for cross-scene classification based on adversarial learning and pareto ranking," *Remote Sens.*, vol. 12, no. 11, 2020, Art. no. 1716.

[20] J. Zhang, J. Liu, L. Shi, B. Pan, and X. Xu, "An open set domain adaptation network based on adversarial learning for remote sensing image scene classification," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2020, pp. 1365–1368.

[21] J. Wang and J. Jiang, "Adversarial learning for zero-shot domain adaptation," in *Proc. Eur. Conf. Comput. Vis.*, 2020, pp. 329–344.

[22] Q. Wang and T. Breckon, "Generalized zero-shot domain adaptation via coupled conditional variational autoencoders," 2020, *arXiv:2008.01214*.

[23] K. C. Peng, Z. Wu, and J. Ernst, "Zero-shot deep domain adaptation," in *Proc. Eur. Conf. Comput. Vis.*, 2018, pp. 764–781.

[24] B. Pradhan, H. Al-Najjar, I. Maher, I. Tsang, and A. Alamri, "Unseen land cover classification from high-resolution orthophotos using integration of zero-shot learning and convolutional neural networks," *Remote Sens.*, vol. 12, no. 10, 2020, Art. no. 1676.

[25] Q. Wang, P. Bu, and T. Breckon, "Unifying unsupervised domain adaptation and zero-shot visual recognition," in *Proc. Int. Joint Conf. Neural Netw.*, Budapest, Hungary, 2019, pp. 1–8.

[26] J. Wang and J. Jiang, "Conditional coupled generative adversarial networks for zero-shot domain adaptation," in *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, 2019, pp. 3374–3383.

[27] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, Oct. 2010.

[28] W. 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.

[29] Z. Wang, J. Jiao, Q. Zeng, and J. Liu, "A three-hierarchy evaluation of polarimetric performance of GF-3, compared with ALOS-2/PALSAR-2 and RADARSAT-2," *Sensors*, vol. 19, no. 7, 2019, Art. no. 1493.

[30] R. Gui, X. Xu, L. Wang, R. Yang, and F. Pu, "A generalized zero-shot learning framework for PolSAR land cover classification," *Remote Sens.*, vol. 10, no. 8, 2018, Art. no. 1307.

[31] W. Sun, P. Li, B. Du, J. Yang, and L. Zhao, "Scatter matrix based domain adaptation for bi-temporal polarimetric SAR images," *Remote Sens.*, vol. 12, no. 4, 2020, Art. no. 658.

[32] H. Dong, X. Xu, R. Yang, and F.-L. Pu, "Component ratio-based distances for cross-source PolSAR image classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 5, pp. 824–828, May 2020.

[33] X. Qin, J. Yang, P. Li, W. Sun, and W. Liu, "A novel relational-based transductive transfer learning method for PolSAR images via time-series clustering," *Remote Sens.*, vol. 11, no. 11, 2019, Art. no. 1358.

[34] W. Zhang, Y. Zhu, and Q. Fu, "Adversarial deep domain adaptation for multi-band SAR images classification," *IEEE Access*, vol. 7, pp. 78571–78583, 2019.

[35] L. Pallotta, A. De Maio, and D. Orlando, "A robust framework for covariance classification in heterogeneous polarimetric SAR images and its application to L-band data," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 1, pp. 104–119, Jan. 2019.

[36] L. Pallotta and D. Orlando, "Polarimetric covariance eigenvalues classification in SAR images," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 5, pp. 746–750, May 2019.

[37] B. Hou, J. Guan, Q. Wu, and L. Jiao, "Semisupervised classification of PolSAR image incorporating labels semantic priors," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 10, pp. 1737–1741, Oct. 2020.

[38] M. Ishii, T. Takenouchi, and M. Sugiyama, "Partially zero-shot domain adaptation from incomplete target data with missing classes," in *Proc. IEEE Winter Conf. Appl. Comput. Vis.*, 2020, pp. 3041–3049.

[39] Q. Song, H. Chen, F. Xu, and T. J. Cui, "EM simulation-aided zero-shot learning for SAR automatic target recognition," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 6, pp. 1092–1096, Jun. 2020,

[40] L. Wang, X. Xu, R. Gui, R. Yang, and F. Pu, "Learning rotation domain deep mutual information using convolutional LSTM for unsupervised PolSAR image classification," *Remote Sens.*, vol. 12, no. 24, 2020, Art. no. 4075.

[41] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, Jul. 2006.

[42] L. Su, J. Shi, P. Zhang, Z. Wang, and M. Gong, "Detecting multiple changes from multi-temporal images by using stacked denosing autoencoder based change vector analysis," in *Proc. Int. Joint Conf. Neural Netw.*, 2016, pp. 1269–1276.

[43] P. Vincent, H. Larochelle, I. Lajoie, Y. Bengio, and P. Manzagol, "Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion," *J. Mach. Learn. Res.*, vol. 11, no. 12, pp. 3371–3408, Oct. 2010.

[44] H. Parikh, S. Patel, and V. Patel, "Classification of SAR and PolSAR images using deep learning: A review," *Int. J. Image Data Fusion*, vol. 11, no. 1, pp. 1–32, Aug. 2019.

[45] L. Liu, Y. Wang, J. Peng, L. Zhang, B. Zhang, and Y. Cao, "Latent relationship guided stacked sparse autoencoder for hyperspectral imagery classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 5, pp. 3711–3725, May 2020.

[46] T. Gadhiy, S. Tangirala, and A. Roy, "Stacked autoencoder based feature extraction and superpixel generation for multifrequency PolSAR image classification," in *Proc. Int. Conf. Pattern Recognit. Mach. Intell.*, 2019, pp. 331–339.
[47] W. Yang, X. Yin, H. Song, Y. Liu, and X. Xu, "Extraction of built-up areas from fully polarimetric SAR imagery via PU learning," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 4, pp. 1207–1216, Dec. 2013.
[48] R. Gui, X. Xu, L. Wang, R. Yang, and F. Pu, "Eigenvalue statistical components-based PU-learning for PolSAR built-up areas extraction and cross-domain analysis," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 3192–3203, 2020.
[49] Y. Xian, Z. Akata, G. Sharma, Q. Nguyen, M. Hein, and B. Schiele, "Latent embeddings for zero-shot classification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 69–77.
[50] Y. Ye, Y. He, T. Pan, J. Li, and H. T. Shen, "Alleviating domain shift via discriminative learning for generalized zero-shot learning," *IEEE Trans. Multimedia*, to be published, doi: 10.1109/TMM.2021.3063616.
[51] J.-S. Lee, M.-R. Grunes, E. Pottier, and L. Ferro-Famil, "Unsupervised terrain classiflcation preserving polarimetric scattering characteristics," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 4, pp. 722–731, Apr. 2004.

Rong Gui received the M.E. and Ph.D. degrees in signal and information processing from Wuhan University, Wuhan, China, in 2016 and 2021, respectively.

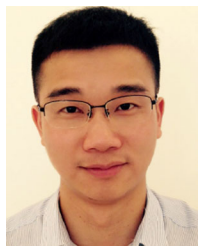
She is currently a Lecturer with the Department of Geomatics and Remote Sensing, School of Geosciences and Info-Physics, Central South University, Changsha, China. Her research interests include synthetic aperture radar imagery processing, 3-D laser measurement technology, and pattern recognition.

Xin Xu received the B.S. degree in electronic engineering, the M.S. degree in signal and information processing, and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 1989, 1996, and 2003, respectively.

He is currently a Professor with the School of Electronic Information, Wuhan University. His research interests include image understanding and statistical signal processing with applications to remote sensing.

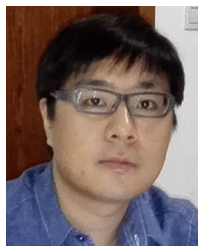
Rui Yang received the B.S. degree in electronic engineering in 2017 from Wuhan University, Wuhan, China, where he is currently working toward the Ph.D. degree in signal and information processing.

His research interests include machine learning and computer vision, especially applications on remote sensing imagery processing.



Kailiang Deng received the Ph.D. degree in marine geodesy from Dalian Naval Academy, Dalian, China, in 2011.

He is currently an Engineer with the Naval Institute of Hydrographic Surveying and Charting, Tianjin, China. His research interests include environmental information acquisition, quality control, and applied research.



Jun Hu (Member, IEEE) received the M.Eng. and Ph.D. degrees in geodesy and surveying engineering from Central South University, Changsha, China, in 2008 and 2013, respectively.

From 2013 to 2014, he was a Postdoctoral Fellow with the Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hong Kong. He is currently a Full Professor with the Department of Geomatics and Remote Sensing, School of Geosciences and Info-Physics, Central South University. He has authored more than 60 papers in international peer-reviewed journals. His research interests include mapping multidimensional and high-precision deformations with interferometric synthetic aperture radar under complicated environment and its applications in geophysical fields.