

Review

The Eyes of the Gods: A Survey of Unsupervised Domain Adaptation Methods Based on Remote Sensing Data

Mengqiu Xu , Ming Wu * , Kaixin Chen , Chuang Zhang  and Jun Guo 

School of Artificial Intelligence, Beijing University of Posts and Telecommunications, Beijing 100876, China

* Correspondence: wuming@bupt.edu.cn; Tel.: +86-133-8330-0814

Abstract: With the rapid development of the remote sensing monitoring and computer vision technology, the deep learning method has made a great progress to achieve applications such as earth observation, climate change and even space exploration. However, the model trained on existing data cannot be directly used to handle the new remote sensing data, and labeling the new data is also time-consuming and labor-intensive. Unsupervised Domain Adaptation (UDA) is one of the solutions to the aforementioned problems of labeled data defined as the source domain and unlabeled data as the target domain, i.e., its essential purpose is to obtain a well-trained model and tackle the problem of data distribution discrepancy defined as the domain shift between the source and target domain. There are a lot of reviews that have elaborated on UDA methods based on natural data, but few of these studies take into consideration thorough remote sensing applications and contributions. Thus, in this paper, in order to explore the further progress and development of UDA methods in remote sensing, based on the analysis of the causes of domain shift, a comprehensive review is provided with a fine-grained taxonomy of UDA methods applied for remote sensing data, which includes Generative training, Adversarial training, Self-training and Hybrid training methods, to better assist scholars in understanding remote sensing data and further advance the development of methods. Moreover, remote sensing applications are introduced by a thorough dataset analysis. Meanwhile, we sort out definitions and methodology introductions of partial, open-set and multi-domain UDA, which are more pertinent to real-world remote sensing applications. We can draw the conclusion that UDA methods in the field of remote sensing data are carried out later than those applied in natural images, and due to the domain gap caused by appearance differences, most of methods focus on how to use generative training (GT) methods to improve the model's performance. Finally, we describe the potential deficiencies and further in-depth insights of UDA in the field of remote sensing.

Keywords: unsupervised domain adaptation; remote sensing data; a survey; deep learning



Citation: Xu, M.; Wu, M.; Chen, K.; Zhang, C.; Guo, J. The Eyes of the Gods: A Survey of Unsupervised Domain Adaptation Methods Based on Remote Sensing Data. *Remote Sens.* **2022**, *14*, 4380. <https://doi.org/10.3390/rs14174380>

Academic Editor: Emanuele Frontoni

Received: 25 July 2022

Accepted: 30 August 2022

Published: 3 September 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Earth observation, climate change and even space exploration have been prominent human concerns in recent years, since they have an impact on human life and production [1–3]. As satellite and imaging technology advance, we are able to collect vast amounts of data from multiple satellites for a variety of uses and detection methods. In general, accurate human annotations for cognitive remote sensing data must rely on expert knowledge, which comes at a significant labor cost. Deep learning algorithms [4–8] have made enormous strides in this area, but the majority of them are data-driven, meaning that the only way to train models for higher performance is with labeled data. However, the present conundrum is that increasing data and annotation capabilities lead to a number of conflicts and contradictions, leaving a lot of data unlabeled and making it challenging to implement practical applications. In order to address the aforementioned issues, Unsupervised Domain Adaptation (UDA) research [9–12] based on deep learning techniques aims to improve the performance of the models learned by labeled data defined as the source domain and inferred in unlabeled data defined as the target domain.

The performance of models noticeably deteriorates when it is directly trained on labeled data and tested on new data, since a significant discrepancy of data distribution exists, which is called the domain shift or domain gap in the unsupervised domain adaptation methods that may be caused by different devices, data modalities, detection areas, seasons or others factors. Thus, the goal of UDA is to distill data and knowledge of different domains to mitigate the effect of domain shift. Our discussion of the influencing factors of domain shift aiming to remote sensing data is summarized into three categories: those related to data acquisition and imaging, those related to tasks and annotations and other factors. In order to illustrate these categories, we provide image examples based on various remote sensing datasets and tasks.

Transferring appearances of data from other domains to be with a similar look [13–20] or a standard style [21,22] is the simplest way to accomplish the goal of UDA, and then a pre-trained model may perform well as a result. We collectively refer to these kind of methods acting on representation as Generative training methods, abbreviated as GT methods. However, the majority of these systems rely on generative methods that sometimes are unstable, making it difficult to guarantee semantic consistency. Thus, adversarial training (AT) and Self-training (ST) are two brand-new learning methodologies to further solve semantic consistency issues. Numerous methods based on Adversarial training (AT) try to identify multi-level domain-invariant features between different domains through calculating distances to narrow margin distribution or joint distribution [23–26], or designing a domain discriminator to achieve domain confusion [27–33]. However, it is difficult for these methods to align same category across domains so that may cause negative transfer on the target domain due to excessive learning of the source domain knowledge. Additionally, Self-training (ST) methods [34–36] can be utilized to achieve unsupervised domain adaptation, such as using pseudo labels generated from a model trained on the source domain, which are used to train or fine-tune a target model using various training approaches. However, pseudo labels may be imprecise due to the lack of class recognition for hard examples. Because there are benefits and drawbacks to each above-mentioned individual training mode, the current trend in unsupervised domain adaptation is to combine them to iterate continuously on the performance of the target domain model [37,38], which are defined as Hybrid training (HT) methods, but these approaches are limited because most of them are designed with multiple stages, causing parameter growth and slow speed.

In real-life or industrial scenarios, unsupervised domain adaptation methods may face more and more complex problems, such as inconsistency in the category space between source and target domain due to the emergence of unknown categories [39–42], multi-domain migration [43–46] due to data collection from different devices, the impediment of weakly supervised learning [47,48] due to rough annotation of source domain data and feature extraction problems [49,50] caused by spectral and temporal shift and so on. Therefore, based on the above problems, the standard UDA methods have some certain extensions, such as Open-Set Domain Adaptation (OSDA) and Partial Domain Adaptation (PDA) under the inconsistent setting of label space and Multi-Domain Adaptation (MDA) for migration between multiple domains.

Despite the fact that numerous publications [9,10,51–53] have already concentrated on unsupervised domain adaptation on real scenes using natural imaging datasets and several works [11,12,54–56] combined deep learning techniques with data from remote sensing area, these works lack a clear description of the domain gap and suffer from the drawbacks of a systematic approach. Besides, there has not been much conversation centered on remote sensing data and specific questions in the field of remote sensing. Therefore, in this paper, we present a comprehensive assessment of unsupervised domain adaptation approaches and applied tasks based on remote sensing data from various satellites or aircraft.

The following are the main contributions:

- This paper clarifies and reviews the idea of unsupervised domain adaptation in remote sensing area. In a nutshell, this paper provides an overview of unsupervised domain adaptation methods divided into four categories: (1) Generative training

methods, (2) Adversarial training methods, (3) Self-training methods, (4) Hybrid training methods. This paper also explores the benefits and potential drawbacks of various training methods by contrasting the statistics of experimental results.

- This paper elaborates on the domain shift based on a thorough review of large remote sensing datasets used in unsupervised domain adaptation research and analyzes its influencing factors, which vary from factors of data acquisition and imaging to factors of tasks and annotation.
- Remote sensing applications with some case studies using unsupervised domain adaptation in remote sensing data are introduced in this paper. Besides, we also concentrate on unsupervised domain adaptation methods based on practical dilemma with remote sensing data, such as multi-domains, partial and open set issues, including task definition, and solution approaches.
- The possible hazards and future development directions of unsupervised domain adaptation methods in remote sensing images are also analyzed in depth through a comparison of methods for natural images.

2. Overview

2.1. Notations and Definitions

Some fundamental concepts relating to data distribution and task function in unsupervised domain adaptation based on remote sensing data should be defined. For deep learning models based on the data domain D , based on task T with the label space Y , the following components need to be formulated: the feature space X with the marginal probability distribution $P(X)$ and a conditional probability distribution $P(Y|X)$. Considering that remote sensing tasks are greatly affected by resolution and detection channels, we have provided definitions for terms such as resolution τ and number of detection bands B .

Aiming for descriptions and definitions of the source domain D_s , we labeled instances $\{x_i^s, y_i^s\}_{i=1}^{n_s}$ with resolution factor τ_s in source domain data X_s and label space Y_s , where $\{x_i^s\} \in \mathbb{R}^{B_s \times H_s \times W_s}$ and y_i^s , depending on the task definition. For example, for a segmentation task, y_i^s is described as $y_i^s \in \{0, 1\}^{b_s \times h_s \times w_s}$, where b_s is the number of classes and h_s and w_s represent the height and width of the image. Moreover, aiming to descriptions and definitions of target domain D_t , we unlabeled instances $\{x_i^t\}_{i=1}^{n_t}$ with resolution factor τ_t in target domain data X_t and label space Y_t , where $x_i^t \in \mathbb{R}^{B_t \times H_t \times W_t}$, where B_t , H_t and W_t represent the number of channels, height and width of the image. It is important to keep in mind that in contrast to natural images B_s or B_t may be more than three and not equal, since many bands are frequently detected at the same time in the field of remote sensing, such as with multi-spectral and hyper-spectral satellites.

The goal of unsupervised domain adaptation is to find a function $F : X \rightarrow Y$, which is obtained by using the source domain data with labels and transferred to classify the target domain with a good performance. Generally, standard unsupervised domain adaptation refers to single-source and single-target domain adaptation when they are shared same label space, as shown in Figure 1a. Besides, in the face of more complex practical scenarios, for example, unknown category or multi-domain transfer is common challenges in realistic applications, we also provide the definition of the above issues as follows.

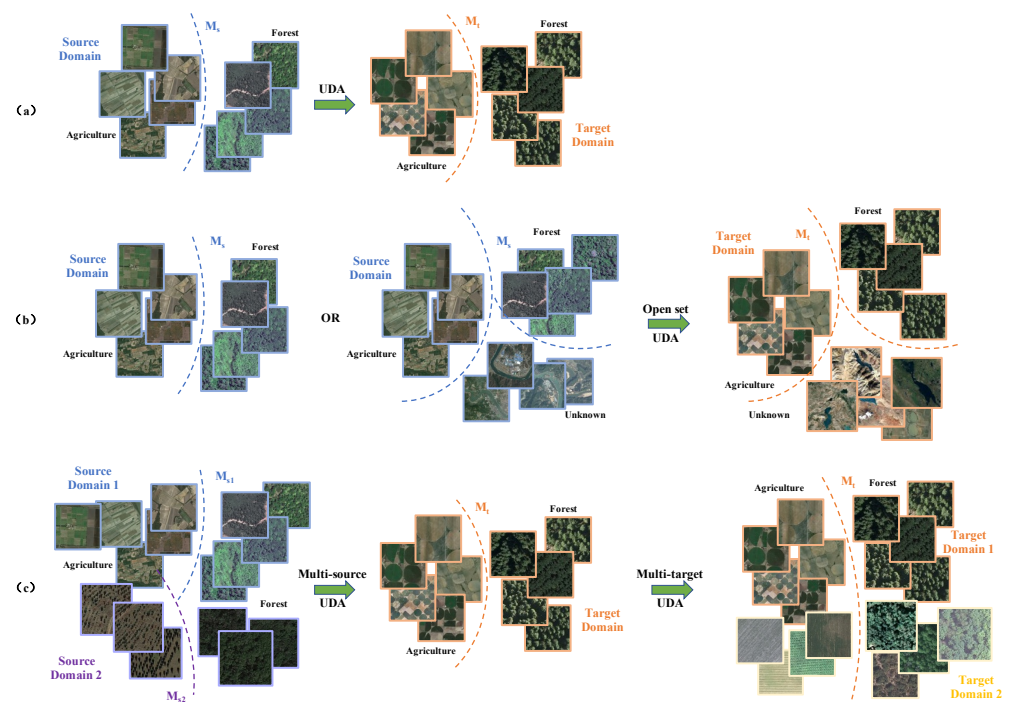


Figure 1. Visualization samples of different definition and setting of unsupervised domain adaptation based on AID, NWPU and PatternNet datasets. (a) Standard unsupervised domain adaptation of a single-source and target domain, which share the same label space. (b) Open-set unsupervised domain adaptation with two different settings, where unknown instances are seen in the source domain and unknown classes that are not the same between two domains and another are unseen. (c) Multi-domain unsupervised domain adaptation divided into multi-source or multi-target domain adaptation.

Closed, partial or open set unsupervised domain adaptation.

Most of the current unsupervised domain adaptation methods are oriented to the same label space, i.e., $Y_s = Y_t$. Source and target domain data share the same semantic category and the unknown classes do not exist during model computation and inference process. However, in the wild, it is common to be faced with the situation that source domain and target domain do not share the same label space; thus, Open Set Domain Adaptation (OSDA) and Partial Domain Adaptation (PDA) came into being. In OSDA settings, the difficulty of OSDA is to isolate unknown categories to avoid misjudgment. In PDA methods, the classes of the target domain constitute a subset of the classes of source domain, that is to say $Y_t \subset Y_s$, and the key of PDA is to avoid negative transfer due to the mismatch classes through identifying the source instances that belong to the same class in the target domain. The classes of target domain contain certain classes and unknown classes that do not exist in the source domain, that is $Y_s \subset Y_t$; the source domain and target domain only share some common classes and they have private classes in their own domain, that is $Y_s \cap Y_t = Y_c$, $Y_s = Y_{sp} \cup Y_c$, $Y_t = Y_{tp} \cup Y_c$, where Y_c , Y_{sp} , Y_{tp} represent the common label space, private source label space and private target label space, respectively. Methods are often divided into two categories according to whether unknown classes are unseen during an unsupervised domain adaptation training process in source domain following OSDA [41] and OSBP [42]; the sample based on the AID and NWPU datasets is as shown in Figure 1b.

Single-domain or multi-domains unsupervised domain adaptation.

In reality, it is possible to be faced with a multi-source domain or multi-target domain to improve the generalization performance of the model and to be applied broadly. The key problem of multi-domain adaptation is mitigating the domain shift between sub-source or sub-target domains. One direct approach is to select samples similar to target domain data

rather than giving up other samples. The sample based on the AID and NWPU datasets is shown in Figure 1c.

2.2. Remote Sensing Datasets and Tasks

Through extensive research of unsupervised domain adaptation in remote sensing data, we selected basic 30 datasets and divided them into 5 categories according to different tasks, namely Regression, Classification, Detection, Segmentation and Generation tasks, with or without class-aware annotations. The dataset descriptions and related works are shown in Table 1. In this paper, we only focus on remote sensing imagery data applied by unsupervised domain adaptation methods containing optical images (Opt), multi-spectral images (MSI) or hyper-spectral images (HSI) and Synthetic Aperture Radar (SAR) images. Because the generation task is not class aware and the regression task contains continuous ground-truth, the Types in the description are marked with ‘-’ in Table 1. From Table 1, we can draw the conclusion that it has a lot of variance across different datasets, although below the same task setting or same detection band, which is also the root cause of the domain shift of data distribution. Besides, different types of remote sensing data are suitable for different tasks; for example, SAR imagery data are more applicable for scene classification or an object detection task, although a few SAR datasets are applied to segmentation task due to the difficulty of image regression with optical imagery. Some samples of classification and segmentation task are shown in Figure 2.

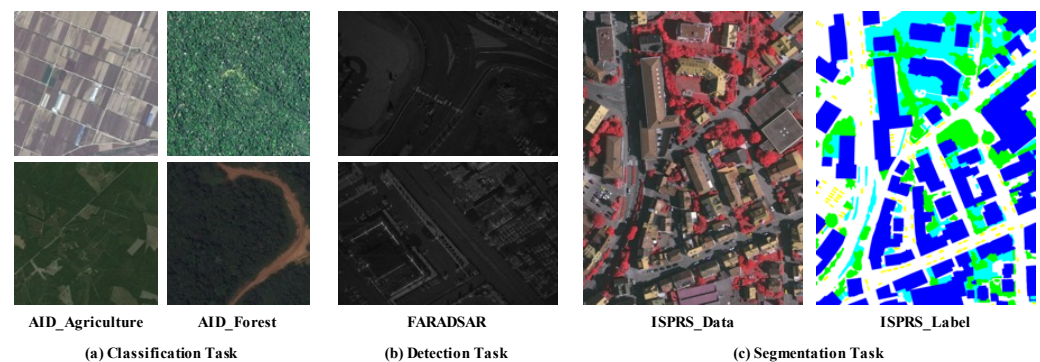


Figure 2. Some samples of unsupervised domain adaptation based on remote sensing data. (a) Samples of the classification task based on the optical AID datasets. (b) Samples of the detection task based on the FARADSAR dataset. (c) Samples datasets of segmentation task, which are the data and labels of ISPRS Vaihingen, respectively.

From the perspective of remote sensing applications [57], unsupervised domain adaptation methods based on deep learning methods are also widely used in scene recognition, object detection, land-cover classification, change detection and 3D reconstruction and other aspects, which have achieved remarkable results to tackle common problems faced by whole humankind, such as climate change, impact of human activities and so on. The following reviews different tasks and shows some application miniature, including Local Climate Zones (LCZ) classification, Vehicle Detection, Change Detection and Dehazing or Decloud, based on unsupervised domain adaptation methods used in the remote sensing circumstance.

Table 1. Some remote sensing datasets used in unsupervised domain adaptation (UDA).

Tasks	Dataset	Description							Works
		Bands	Types	Device	Size-I	Size-D	Resolution	Region	
Regression	SARptical [58]	SAR-2 Opt-3	-	TerraSAR-X UltraCAM	112 × 112 112 × 112	20,216	1 m 0.2 m	Berlin	
Classification	NWPU [59]	Opt-3	45	Google Earth	256 × 256	31,500	0.2–30 m	-	[60–63]
	PatternNet [64]	Opt-3	38	Google Map	256 × 256	800	0.062–4.693 m	-	
	AID [65]	Opt-3	30	Google Earth	600 × 600	10,000	0.5–8 m	-	
	Merced [66]	Opt-3	21	Map	256 × 256	2100	0.3 m	-	
	Eurosat [67,68]	Opt-3	10	Sentinel-2	64 × 64	27,000	0.2–30 m	-	[70]
		MSI-13	10	Sentinel-2	64 × 64	27,000	0.2–30 m	-	
	MSTAR [69]	SAR-2	10	SAR Sensors	128 × 128	17,658	0.3 m	-	
	So2Sat LCZ42 [71]	SAR-8	17	Sentinel-1	32 × 32	400,673	10 m	52 cities	
		MSI-10		Sentinel-2	32 × 32		10 m		
Detection	DOTA [72]	Opt-3	15	Google Earth, GF-2, JL-1	800 × 800–4000 × 4050	2806	-	-	[73]
	Dior [74]	Opt-3	20	Google Earth	800 × 800	23,463	0.5–30 m	-	
	Optical-SAR [75]	Opt-1	5	Google Earth	192 × 192	10,000	0.3 m	Visakhapatnam	
		SAR-1		TerraSAR-X					
	Simulation Data [75]	SAR-2	6	3D-CAD	-	10,800	-	-	
	Measured DataSet [75]	SAR-2	5	Sentinel-1,TerraSAR-X	192 × 192	17,500	5/1 m	-	
	FARADSAR [76]	SAR-1	10	Radar	1300 × 580–1700 × 1850	106	0.1 m	The University of New Mexico	[37]
	miniSAR [77]	SAR-1	10	Radar	1638 × 2510	9	0.1 m	Kirtland Air Force Base	
Segmentation	Skyscape [78]	Opt-3	31	Camera System	5616 × 3744	16	0.13 m	Munich	[81]
	LoveDA [79]	Opt-3	7	Spaceborne	1024 × 1024	5987	0.3 m	NanJing, ChangZhou, WuHan	
	ISPRS Vaihingen [80]	Opt-3	5	-	2500 × 2000	32	0.09 m	Vaihingen	
	ISPRS Potsdam [80]	Opt-3	5	-	6000 × 6000	38	0.05 m	Potsdam	
	Beijing Dataset [81]	Opt-3	4	DigitalGlobe, SpaceView	1800 × 800	202	0.3 m	Beijing	
	Massachusetts [82]	Opt-3	3	-	1500 × 1500	1171	1 m	Massachusetts	
	SpaceNet [83]	SAR-5	2	Aerial Sensor	406–439	6000	0.5 m	Rotterdam	
		Opt-3	2	WorldView-2	406–439	6000	0.5 m	Rotterdam	
	DeepGlobe [84]	Opt-3	2	DigitalGlobe, Vivid	1024 × 1024	8970	0.5 m	Thailand, Indonesia,	
	CHN6-CUG [85]	Opt-3	2	Google Earth	512 × 512	4511	0.5 m	India, 6 cities	

Table 1. Cont.

Tasks	Dataset	Description							Works
		Bands	Types	Device	Size-I	Size-D	Resolution	Region	
	Indian Pines [86]	HSI-224	16	AVRIS Sensor	145 × 145	10,249	20 m	North-Western Indiana	
	Salinas [87]	HSI-224	16	AVRIS Sensor	512 × 217	54,129	3.7 m	Salinas Valley	
		HSI-224	16	AVRIS Sensor	86 × 83	5348	-	Salinas Valley	
	Botswana [88]	HSI-145	14	NASA EO-1	1476 × 256	3284	30 m	Okavango Delta	[81]
		HSI-145	14	EO-1 Satellite	1476 × 256	2494	30 m	Okavango Delta	
	Kennedy Space Center [89]	HSI-176	13	Spectrometer	512 × 614	1	18 m	Kennedy	[90]
	Washington DC MALL [91]	HSI-191	7	Sensor	1280 × 307	1	-	Washington	[92]
Generation	RICE [93]	Opt-3	2	Google Earth	512 × 512	1000	-	-	
		Opt-3	4	Landsat 8 OLI/TIRS	512 × 512	450 sets	-	-	
	SEN12MS-CR [94,95]	SAR-2		Sentinel-1	256 × 256		10 m		
		MSI-13	-	Sentinel-1	256 × 256	472,563	10 m	-	
		MSI-13		Sentinel-2	256 × 256		10 m		

2.2.1. Classification Task

A classification task based on image level is the classic research both in the field of remote sensing and deep learning that is the basis of detection and segmentation task. With the development of data acquisition and imaging technology, taking the aerial scene classification task as an example, many mature UDA techniques [60–63], varying from single-domain transfer to multi-domains, have been developed based on diverse datasets, such as NWPU [59], PatternNet [64], AID [65] and Merced [66] datasets, as shown in Table 1. The focus of these works is how to find domain-invariant features to ensure semantic consistency. Despite to common classification tasks, such as land type classification, ship classification, etc., in the field of remote sensing, it is worth following with interest the **Local Climate Zones (LCZ) classification** [71], which was originally proposed for urban heat island studies. As mentioned by [96], transfer of an LCZ task may cause the degradation of model performance due to a domain shift and encourages most advanced UDA methods applied in this area. Subsequently, CPDA [97] presented a circled similarity propagation-based domain adaptation method, while [98] used the co-training approach with self-paced learning to achieve good performance of the LCZ task.

2.2.2. Detection Task

The detection task not only requires the model to recognize the semantics of the multi-class object, but also need to mark the location with coordinate information using a bounding box. Moreover, it has more subdivided applications in the field of remote sensing, such as ship or vehicle detection [99–102], used for national defense security, and oil palm tree detection [103–106], used for economic production, and can even give cross-temporal change detection to make people understand the world more macroscopically. We take **Vehicle Detection** [99–102] based on unsupervised domain adaptation methods as one case study. Even though the vehicle detection task has become a hot topic and application in remote sensing imagery, model performance trained on a certain amount of labeled data based on deep learning methods [6,107] would degrade when faced with new data collected from new areas or satellites [100]. References [100,101] used distance-based and hierarchical-level adversarial training strategy to further extract more discriminative features, respectively. Besides, many works [99,102] also focus on other factors that cause model performance to degrade in the vehicle detection task; for example, reference [102] focused on the discrepancy between daytime and nighttime images and reference [99] transferred the model trained on satellite images to Unmanned Aerial Vehicle (UAV) images.

Besides, in the field of remote sensing, **Change Detection** [108–113] is worthy to be mentioned, since changes usually occur in earth observation as time passes, inevitably leading to differences in geographical appearances. Faced with changes, domain adaptation methods can provide an approach to detection area discrepancy. For example, DSDANet [108] is the first method in which unsupervised domain adaptation is introduced to change detection, and reference [113] proposed an unsupervised domain adaptation method based on CycleGAN [114] to alleviate the domain shift on the deforestation detection application. Based on this work, many works explored this inspiration to multi-source change detection [109] or specific applications, such as monitoring the deforestation or growth of forests [110–112]. For building change detection task, FODA [115] addresses the pseudo-change problem through feature space alignment and adds more effective and discriminative feature extracted by output space by reducing the difference between prediction and the ground truth.

2.2.3. Segmentation Task

Compared with the classification task, the segmentation task is supposed to achieve the meticulous recognition of each pixel and it shows a huge application value and potential. In various segmentation tasks, including road extraction, building extraction, agriculture,

vegetable segmentation and so on, **Land Use or Land Cover Classification** [49,81,116–118] based on pixel-level is a common but hot topic. Taking road extraction as a case study, some methods [32,119,120] utilize adversarial training methods to narrow the gap between source and target domain; in particular, reference [31] converted the deep features extracted by the convolutional network to 2D features, such as curves, to more effectively ensure domain-invariant features. Recently, a lot of works [121,122] attempted to implement road extraction on target domain with a two-stage network or the hybrid training mode. In addition, some works [79,123] paid attention to the discrepancy brought by different time or space to remote sensing data, such as the differences in urban or suburban areas due to the degree of economic development. Furthermore, we review the adaptation development of the Land Use Classification task based on the ISPRS Potsdam and Vaihingen datasets in detail in Section 5, in which the performances using different training mode methods are listed and compared.

2.2.4. Generation Task

With the rapid development of image translation and multi-modal technology, especially the proposal of Generative Adversarial Network (GAN) methods, generative tasks can not only achieve style transfer under the condition of invariant semantics, but they can also synthesize, generate or create new images according to potential learning principles under the guidance of semantics. However, in the actual data acquisition process, the Generation task is often faced with the inability to obtain paired data, and the existing data and the generated data are quite different, so they cannot obtain a good generation performance. We take the **Dehazing or Decloud** [124–126] task as an example to introduce how unsupervised domain adaptation methods are developed and applied in reality. Due to fog, haze and other bad weather conditions, the contrast and color fidelity of remote sensing images degrades, which makes it is hard to obtain a model with good generalization. The most common solution is using image translation methods to achieve the image dehaze or decloud process [95,127], such as CycleGAN or conditional GAN. However, in the face of complex and diverse weather conditions and data, sometimes, image translation makes it difficult to ensure semantic consistency. Thus, domain adaptation methods provide a new approach [124] to solve hazy or foggy remote sensing data recognition problems, where clear images can be defined as source domain and hazy or foggy data as target domain, although without labels. For example, SkyGAN [126] proposed a domain-aware hazy-to-hyperspectral (H2H) module based on cycle-consistency and a domain classifier to achieve image dehazing, while TA^3N [125] achieves degraded remote sensing classification through adversarial training of conjunction with effective image-level and region-level features.

2.3. The Key Problems: Domain Shift

It is easy to reach a consensus on that the core problem that the unsupervised domain adaptation needs to solve is how to narrow the domain shift between the source and target domain. However, it is unclear what factors cause the domain shift that may involve completely diverse color distribution, texture characteristics and contextual information. In accordance with the data generation process and labeling process, we mainly divide the influencing factors into three categories as follows. We provide some samples based on remote sensing data, as shown in Figure 3, for illustration.

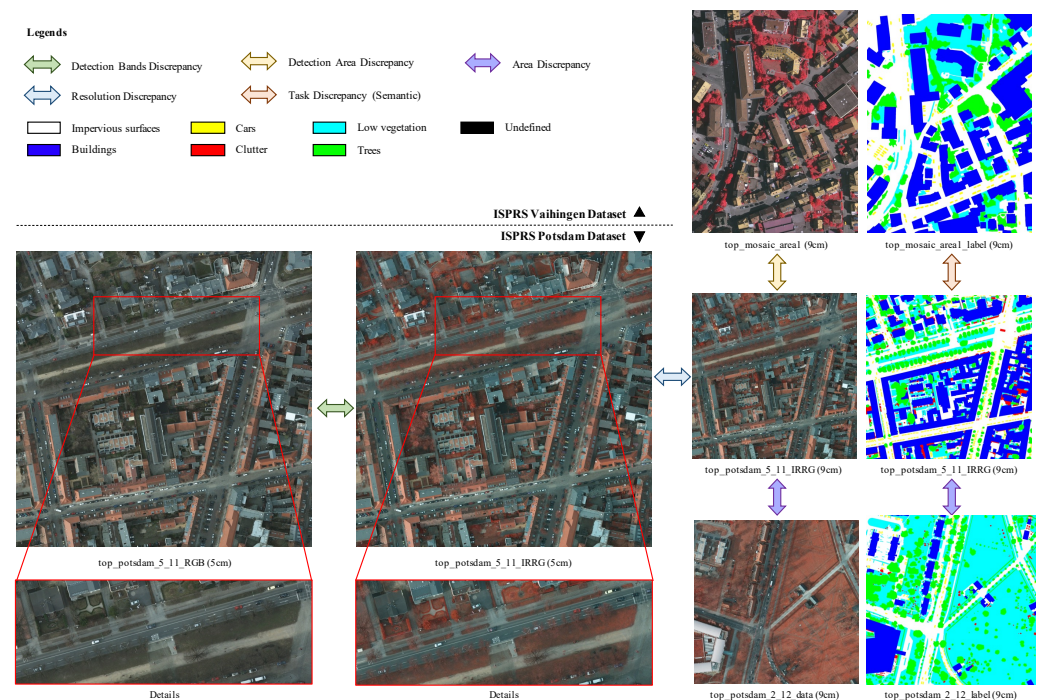


Figure 3. Some samples of an unsupervised domain shift based on the ISPRS Vaihingen and Potsdam dataset. Arrows denote differentiating factors between images and rectangles denote different semantic categories in label masks.

Factors related to data acquisition and imaging. The difference between remote sensing image data acquisition and imaging mainly affects the data distribution, that is, it involves source and target domain data X_s, X_t under the marginal probability distribution $P(X_s), P(X_t)$ and a conditional probability distribution $P(Y_s|X_s), P(Y_t|X_t)$.

- **Detection Areas.** There are huge differences between areas or countries due to the discrepancy of economy level and human activity influence. For example, roads in urban and suburban areas have different characteristics, such as color, connectivity and contraposition with background, as mentioned in LoveDA [79]. From the two images pointed at by the purple arrow in Figure 3, we can also make a sense of the diversity between different areas collected from Potsdam through the density of the building marked in blue in the label images.
- **Illumination.** Imagery is captured at different times of the day as a result of the illumination difference, which may increase the instability of the training model. In particular, visible bands are not available in some all-day monitoring missions at night, which causes discrepancies compared with daytime data.
- **Resolution or Ground Sampling Distance (GSD)** [15,31], taking the Potsdam Dataset with 5 cm resolution and 9 cm resolution as an example, as shown in the two images pointed at by the blue arrow in Figure 3. High resolution and wide coverage may cause images containing redundant information or noises. Reference [60] uses five diverse remote sensing datasets to discuss characteristics through transfer learning, the conclusion of which indicates that multi-resolution is helpful to affect generic representation.
- **Devices and detection bands** [128]. This is quite different between wavelength, type and the number of bands. The two images and details pointed at by the green arrow in Figure 3 show appearance diversities, although they are selected from the same area and satellite but different bands, where one consists of RGB bands and another of IRRG bands. Reference [128] noticed the discrepancy caused by various multi-spectral bands and built domain adaptation methods using RGB images to other types of multi-spectral images. As shown in Figure 4a, we also display the pixel value statistics

for various channels across the two datasets. It is clear that the two datasets' pixel value distributions are very dissimilar. Additionally, the IR-Band differs significantly from the visible light channel after being translated to the 0–255 range.

- Inconsistency in class distribution. Even if we collect data from different domains for the same category, the proportion of a certain category may be different, as shown in Figure 4b, where the 'Building' class has the highest proportion of pixels in the Vaihingen dataset, while 'Impervious surface' is higher in the Potsdam dataset.

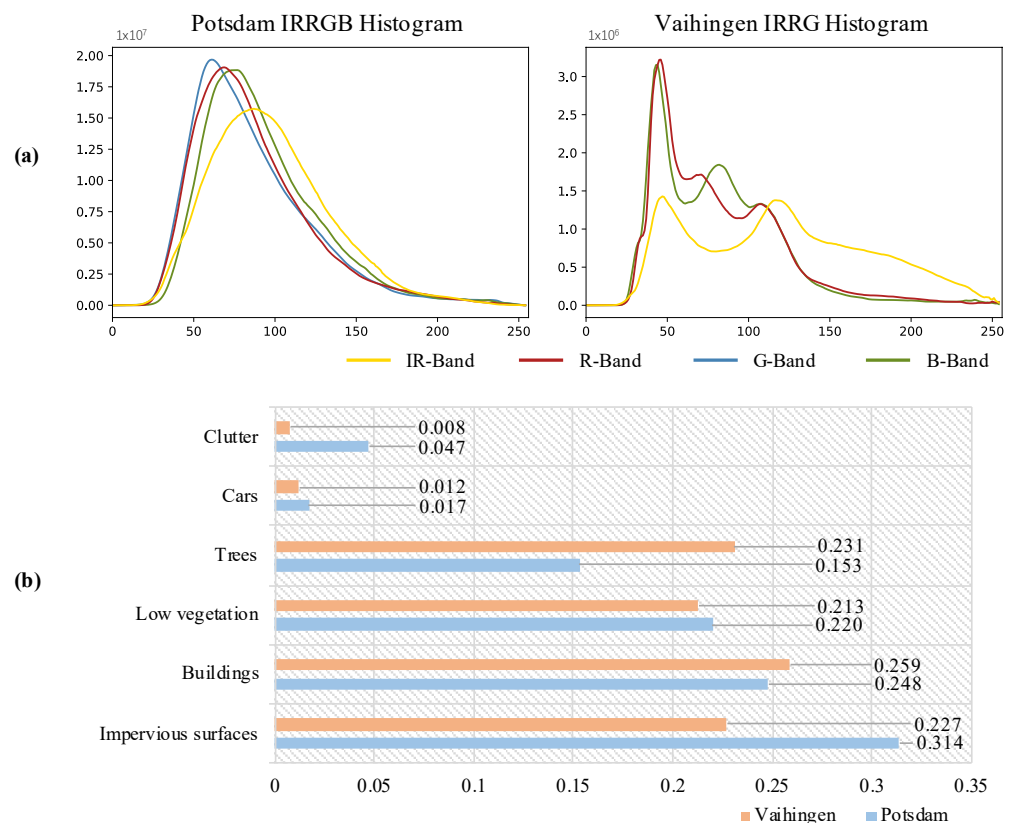


Figure 4. (a) The pixel value statistics of the ISPRS Vaihingen and Potsdam datasets. (b) The proportion of different categories from two datasets. Both of these statistic values indicate large differences between the two datasets, which make data migration difficult.

Factors related to tasks and annotations. Different tasks determine the type and form of annotation, and the distribution of categories is also one of the important factors affecting data distribution, that is, it involves source and target label space Y_s, Y_t . The source and target domain often do not share the same label space so that unseen or unknown class samples may appear in the target domain, that is, the 'Clutter' category does not appear in the Vaihingen image, as shown in Figure 3 and details pointed at by the orange arrow.

Other factors. Different from images in natural scenes that are mainly affected by human activities, such as object detection and autonomous driving data, remote sensing images are easily affected by factors such as solar activity and atmosphere, so accurate detection wavelength data cannot be obtained.

- Atmospheric effects. Sometimes even images collected by the same satellite sensors might have quite different radiometry, which makes them hard to annotate and recognize.
- Position of the sun and satellite observation direction. The imaging quality of Unmanned Aerial Vehicles (UAV) are influenced by position of sun mentioned in [129], such as the incorrect exposures and distortion. The relative position between the sun and the satellite affected the quality of the satellite image in reference [130].

The factors that cause a domain shift need to be discussed in detail according to the setting of the task and the selection of the dataset. For example, building extraction is influenced by the regional differences caused by different climates; however, road extraction is affected by the level of economic development and the impact of topography. Besides, tasks are influenced by seasonal variations as mentioned in SeCo [63]. The predictions for certain tasks, such as land-cover classification, remain the same, even though seasonal factors differ, while others tasks, such as change detection, are easily affected by seasonal factors.

3. Approaches of Domain Adaptation in Remote Sensing

When faced with a new dataset or new circumstances, the fine-tuning [131–133] method is an easy approach to improve the transferability of the model, especially with data-driven algorithms such as deep learning methods. However, generally, it is hard to fine tune due to unavailable annotations for the target domain, and it shows poor performance when there is a much larger domain shift between the source and target domains. Thus, in this section, we mainly provide a systematic overview of existing generative training methods, adversarial training methods, self-training methods and hybrid methods. The methods in this section are basically based on a single-source and target domain, and the multi-domain methods are described in Section 4.

The relevant statistics for UDA used in the natural and remote sensing dataset are shown in Figure 5. These methods take the training mode as the x -axis and the task as the y -axis, and each cluster of methods displayed is sorted by publication year. In addition, we create quantitative histograms from various perspectives on the right and above to compare methods for the two scenarios.

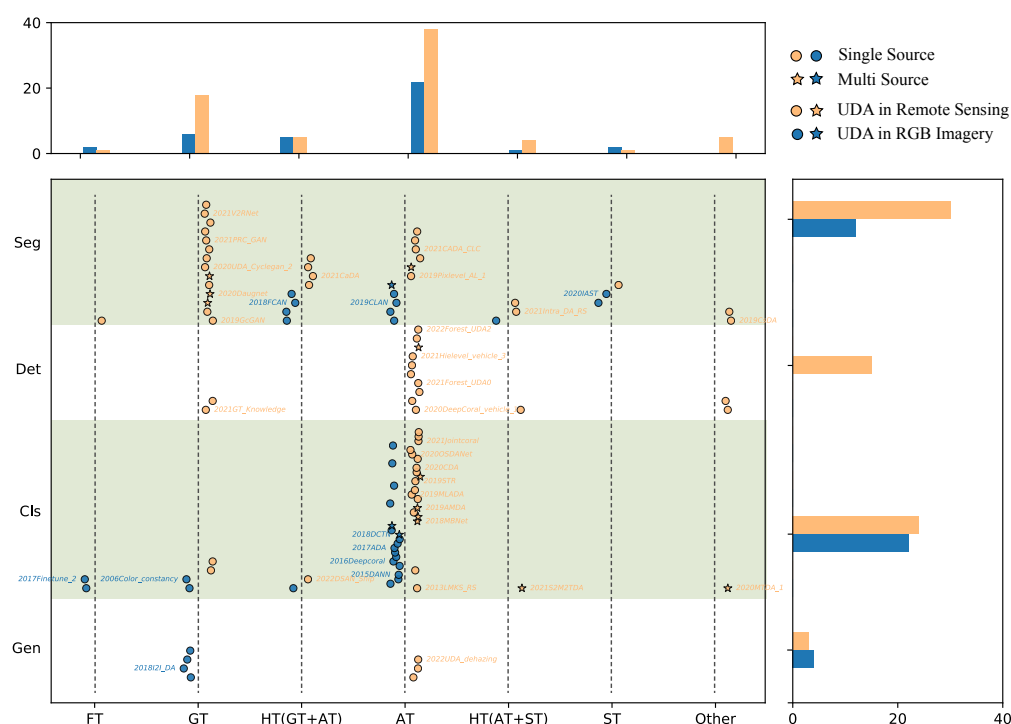


Figure 5. Relevant statistics of unsupervised domain adaptation methods sorted by published year, task and training mode, where geometric shapes in yellow and blue represent methods used in remote sensing and nature scenes, respectively. Circles and stars represent the single-domain and multi-domain methods, respectively. The horizontal axis of the bottom-left figure represents different training modes, where FT, GT, AT, ST and HT are Fine-tune, Generative, Adversarial, Self- and Hybrid training methods, respectively. The vertical axis of the bottom-left figure represents different tasks in the field of remote sensing, where Gen, Cls, Det and Seg represent the Generation, Classification, Detection and Segmentation tasks, respectively.

3.1. Generative Training Methods

Generative training (GT) methods mainly focus on discrepancy of appearance, such as color and texture. The goal of these methods is to generate visually similar images in order to use labels from the source domain to develop a high-performing model that is also adopted in the target domain. Thus, the key issue of generative methods is how to generate fake images under the guidance of semantics and constraints. These methods with formulations are mainly divided. Besides, many methods [134,135] are based on GT methods and combined with other training methods introduced in Section 3.4, such as AT or ST, which have a multi-stage structure, and as a result, they can use the transferred fake image as an intermediate result in other training stages.

- Target-stylized methods. The main purpose of target-stylized methods is to convert the source domain image into a fake image with a similar style to the target domain image so that the semantic consistency of data and labels can be preserved to the greatest extent. The general pipeline is designing a Generator $G_{S \rightarrow T}$ to generate target-stylized images $G_{S \rightarrow T}(s)$, and then utilizing $G_{S \rightarrow T}(s)$ and $L(s)$ to train or fine-tune a model $M = M_s = M_t$. Some representative papers are ColormapGAN or Neural Style Transfer (NST) methods [13,14], GAN methods [15,16,20,116] and so on. Matching methods only consider color distribution matching and do not consider semantics, but migration can be implemented quickly and efficiently when the domain gap is small, such as Graph matching [136], Histogram matching [137], etc.
- Source-stylized or mid-domain transferring methods. Different from target-stylized methods, the biggest advantage of these methods is that they can ensure that the data and labels of the training model are precisely aligned, which can guarantee the accuracy of the semantics. These methods are mainly for transferring target images to source-styled images or to find an intermediate domain between the source and target domain, such as Inverse DA and other multi-domain methods [19,21,22]. The pipeline is similar to the target-stylized pipeline but the data flow is opposite and a Generator defined as $G_{T \rightarrow S}$.

3.1.1. Target-Stylized Methods

Target-stylized methods mainly focus on how to convert the original style of source domain imagery to the style of target domain, which can be divided into three parts as follows. The most direct way to change the style is to adjust the image color representation of different domains to be consistent, as introduced in the first part. However, after a lot of practice, it can be found that the difficulty of Target-stylized methods lies in the consistency of image semantics before and after the transferring process. Thus, the methods that consider cycle-consistency and identify the constraints generally are mostly selected as baselines, and the preliminary migration effect can be achieved as introduced in the second part. In addition, utilizing other information, such as geometry information, can contribute to the semantic invariance.

The first part: Non-GAN methods.

Traditional data augmentation methods are mainly based on matching methods, such as Histogram matching and its variants [137,138], Graph matching [136], color constancy algorithm [139], gray world [140], etc. Although these methods can quickly achieve transferring, they are insufficient when faced with a huge domain shift and new data; besides, it is hard to obtain a good performance when faced with multi-target domains. For example, Randomized Histogram Matching (RHM) [138] matches the histogram of a source-target pair randomly selected from a different domain instead of matching each histogram of the source to the whole target domain data.

ColormapGAN [13] proposes to generate fake training images by learning transferring color features. Different from GAN [141], this paper does not have any convolution or pooling layer in Generator, where the Generator consists of one element-wise matrix multiplication and one matrix addition operations. By training the initial classifier, ColormapGAN and Fine-tuning, classification is achieved.

SemI2I [14] introduced a new data augmentation method by using a style transfer method, such as AdaIN [142], to transfer the style of the target domain data to the training data, whereas in the AdaIN module, σ and μ carry style information of two domains. This method can promise style similarity and semantic consistency at the same time. In this paper, the segmentation network [7,143] firstly trains on original data and then fine-tunes on stylized data.

The second part: GAN methods.

Image-to-image translation applied in domain adaptation [144] source images mainly uses unpaired images, which are similar target images in color or texture or style. An image-to-image translation task based on GAN [141] methods is mainly divided into two parts based on whether data are paired. Since unsupervised domain adaptation does not satisfy the demand that target domain annotation is available, GAN methods based on image translation use unpaired data. It is necessary for source domain data to convert a style-like target domain. From the point of view of data flow, the GAN methods can be divided into single methods [141,145] or cycle methods [114,146]. However, we divide the method from the perspective of unsupervised constraints as follows, because they are related to methods designed for the domain gap.

The adversarial constraint [141] from a basic single Generative Adversarial Network containing a Generator $G_{S \rightarrow T}$ and a Discriminator D_T , with corresponding GAN loss, is as follows, where $\hat{p}_S, \hat{p}_T$ are the discrete distributions sampled from the source and target domain, respectively:

$$L_{adv}(G_{S \rightarrow T}, D_T, \hat{p}_S, \hat{p}_T) = \mathbb{E}_{x_T \sim \hat{p}_T} [\log D_T(x_T)] + \mathbb{E}_{x_S \sim \hat{p}_S} [\log(1 - D_S(G_{S \rightarrow T}(x_S)))] \quad (1)$$

When the data flow is bi-directional or multi-directional, GAN loss is supposed to be calculated twice or more times according to the number of data flow; for example, in CycleGAN, GAN loss should be defined as $L_{adv} = L_{adv}(G_{S \rightarrow T}) + L_{adv}(G_{T \rightarrow S})$.

Cycle-consistency constraint [114]. During the translation process, for a source domain sample, after target-stylized and source-stylized transfer, the style of sample is not changed, and vice versa. Cycle-consistency loss is defined as follows:

$$L_{cyc}(G_{S \rightarrow T}, G_{T \rightarrow S}) = \mathbb{E}_{x_S \sim \hat{p}_S} Dis(G_{T \rightarrow S}(G_{S \rightarrow T}(x_S)) - x_S) + \mathbb{E}_{x_T \sim \hat{p}_T} Dis(G_{S \rightarrow T}(G_{T \rightarrow S}(x_T)) - x_T) \quad (2)$$

where the function Dis can be defined according to different task, such as L_1 distance, L_2 distance and KL divergence. Taking L_1 distance as an example, the above formula can be described as:

$$L_{cyc}(G_{S \rightarrow T}, G_{T \rightarrow S}) = \mathbb{E}_{x_S \sim \hat{p}_S} \|(G_{T \rightarrow S}(G_{S \rightarrow T}(x_S)) - x_S)\|_1 + \mathbb{E}_{x_T \sim \hat{p}_T} \|(G_{S \rightarrow T}(G_{T \rightarrow S}(x_T)) - x_T)\|_1 \quad (3)$$

Identity constraint [147]. Although the style of the source domain sample is changed, the semantics of the sample are maintained, as described as follows, where Dis is defined as above:

$$L_{idt}(G_{S \rightarrow T}, G_{T \rightarrow S}) = \mathbb{E}_{x_S \sim \hat{p}_S} Dis(G_{S \rightarrow T}(x_S) - x_S) + \mathbb{E}_{x_T \sim \hat{p}_T} Dis(G_{T \rightarrow S}(x_T) - x_T) \quad (4)$$

Reference [15] proposed a method for unsupervised domain adaptation in the ISPRS dataset, which contains two steps: the first step is to convert the image of the source domain to the target domain and the second is to fine-tune the model trained on the source domain. Reference [116] firstly trained a base model with encoder-decoder architecture based on SegNet [148]. This paper hoped to achieve adaptation to target domains according to two steps; one is using CycleGAN [114] to achieve translation between the two domains, another is continuously fine-tuning the task network during the translation process.

UST-DG [16] firstly developed DualGAN [149] to achieve unsupervised style transfer to generate target-stylized images for source domain images to alleviate the disadvantage influence of the data shift. Then, a model adapted to the target domain was trained by using pseudo images with labels, obtaining a good performance during the test. In addition, in this paper, what causes the domain shift is discussed through two experiments based on different settings, P(IR-R-G) to V(IR-R-G) and P(R-G-B) to V(IR-R-G).

BiFDANet [20] proposes an unsupervised domain adaptation method based on a bi-direction image-to-image translation method to take full advantage of both domains and overcome poor performance due to unidirectional domain adaptation, while this method uses Deeplab v3 [8] and ResNet [4] as backbones. To keep semantic consistency, three constraints are designed as follows, which are a cycle-consistency constraint [114], an identity consistency and a semantic consistency inspired by papers [43,150]. For the test, a linear combination is designed to merge the prediction results.

Reference [151] proposed a method for Synthetic Aperture Radar (SAR) ship instance segmentation based on cross-domain transfer learning and a res-pyramid Network, where cross-domain transfer learning contains a sample transfer module and a knowledge transfer module. Knowledge transfer of the data is used to train the instance segmentation model, and then the parameters are cropped to retain the relevant parameters at the feature level.

The third part: combination with other constraints.

GcGAN [17] proposed a method for unsupervised domain mapping, which combines the cycle-consistency constraint [114] and geometry-consistency constraint. UGCNet [18] added geometry information proposed by GcGAN [17] to unsupervised domain adaptation methods to alleviate the challenge of obtaining sufficient training data, which consists of a Cross-domain Adaptation Network (CAN) and a Geometry-Consistent Segmentation Network (GSN). The geometry constraint is as follows. It can be assumed that there exists a geometric transformation function $F_{geo}(x)$ with an inverse function $F_{geo}^{-1}(x)$ that satisfied $F_{geo}(G_{S \rightarrow T}(x)) = G_{S \rightarrow T}(F_{geo}(x))$, where $x \in \{x_s, x_t\}$ and $x = F_{geo}^{-1}(F_{geo}(x))$, so that it can be achieve geometry consistency during domain adaptation. The detailed definition is as follows using L1 distance loss marked as $\|\cdot\|_1$:

$$L_{geo}(F_{geo}, G_{S \rightarrow T}) = \mathbb{E}_{x_s \sim \hat{p}_S} \|G_{S \rightarrow T}(x) - F_{geo}^{-1}(G_{S \rightarrow T}(F_{geo}(x)))\|_1 + \mathbb{E}_{x_s \sim \hat{p}_S} \|F_{geo}(G_{S \rightarrow T}(x)) - G_{S \rightarrow T}(F_{geo}(x))\|_1 \quad (5)$$

Based on UGCNet [18], V2RNet [120] is proposed by adding a semantic discriminator in style transfer net named SegGAN to unify the source domain semantic structures and target domain image style. Besides, this work also proposes a U-B-Net to achieve road extraction aiming to the slender shape and uneven distribution of the road. Contemporaneous work [117] also used geographic information, although only the rotation function was used, and explored the differences in the transfer performance between images of different modes.

Besides, there are also some constraints that are not introduced into the domain adaptation methods, such as distance constraint from Distance GAN [152], which is not complete for methods to construct a multi-level constraint system. The distance constraint defined as follows assumes that if two samples obeyed the same distribution and if they are mapping to the same space.

$$L_{dis}(G_{S \rightarrow T}) = \mathbb{E}_{x_{Si}, x_{Sj} \sim \hat{p}_S} \left| \frac{1}{\sigma_S} (\|x_{Si} - x_{Sj}\|_1 - \mu_S) - \frac{1}{\sigma_T} (\|G_{S \rightarrow T}(x_{Si}) - G_{S \rightarrow T}(x_{Sj})\|_1 - \mu_T) \right| \quad (6)$$

3.1.2. Source-Stylized or Mid-Domain Methods

Differently, Inverse domain adaptation [19] thinks transferring target data to source domain style can use more comprehensive features of source domain. During the trans-

lation process, it is necessary to preserve the image details and semantic consistency, for which CycleGAN and AdaIN are used in this paper.

Similar to FCAN [134] consisting of two main components, that is, the Appearance Adaptation Networks (AAN) and the Representation Adaptation Networks (RAN), transferring the appearances of different domains to be domain-invariant by searching a mid-source can achieve domain adaptation, but this idea is frequently used in multi-domains adaptation, since the distance between target domain to different sub-source domains is not the same, such as StandardGAN [21] introduced in detail in Section 4.

3.2. Adversarial Training Methods

The goal of Adversarial training (AT) methods is to extract information on feature-level or pixel-level or both levels to reduce the discrepancy between different domains. Some feature-level domain adaptation methods narrow the domain gap between the source and target domain, e.g., aligning second-order statistics [25,153], contrastive domain discrepancy [34], maximum mean discrepancy [24,154] or Wasserstein metrics [26], while other methods use adversarial training. Compared with generative methods, although pixel-level domain adaptation methods focus on every pixel but not generate mid-output results during the transferring process.

3.2.1. Feature-Level Adversarial Training Methods

The first part: Distance-based methods.

There are two main aspects of distance-based methods that are worthy to be explored: one considers at what level or feature the distance can be measured, and the other considers how to measure the distance. Therefore, we explore the above from the three directions of data consistency, label consistency and joint distribution consistency.

How to consider the distance based on consistency of data or features? Maximum Mean Discrepancies (MMD) [24] and Multi-Kernel Maximum Mean Discrepancies (MK-MMD) [154] are firstly introduced into domain adaptation methods, and reference [155] proved the effectiveness of minimizing the distance between source and target domain. Due to the difficulty of MMD for controlling class distribution, paper [156] proposed a method by learning the manifold embedding and migrate domain gap by embedding space based on class-wise MMD that considers category distribution by adding class indicator function. A two-stage Deep Domain Adaptation (TDDA) [157] method is proposed to achieve hyper-spectral image classification based on three criteria, including MMD and margin-based loss in the first stage, and pairwise loss in the second stage. Besides, a Deep Siamese Domain Adaptation Convolutional Neural Network (CNN) called DSDANet [108] used MK-MMD [154] of calculating distance between spatial-spectral features extracted by a siamese CNN from different domains; this paper claims that it is the first time domain adaptation method used on change detection application.

For different scenarios and applications, a lot of work has carried out function optimization and improvements on the MMD or MK-MMD function. For example, the Class-wise Distribution Adaptation network named CDA [158] proposes the probability Maximum Mean Discrepancies (PMMD) in conjunction with adversarial adaptation to obtain an unsupervised classifier for hyper-spectral remote sensing data, where PMMD uses the probability predictions of target data instead of distribution embedding to Reproducing Kernel Hilbert Spaces (RKHS) in MMD [24] when estimating the means of each class during the adaptation process. In addition, paper [159] used style transferring and Local Maximum Mean Discrepancies (LMMD) proposed by DSAN [160] to achieve fine-grained ship classification. Compared with MMD, LMMD is further supposed to capture fine-grained feature by considering weights between different samples.

In addition to MMD, many methods [25,153] also use other metrics to measure the distance between two domain features. For example, inspired by Coral [153], Deep Coral [25] constructs a loss function with a non-linear transformation named the CORAL Loss based on covariance matrices to minimize the correlations between different domains.

Reference [100] applied correlation alignment and adversarial domain adaptation to improve the accuracy of target domain vehicle detection, and in order to learn semantic features, it utilizes reconstruction loss. Reference [161] introduced domain-level and class-level correlation alignments (CORAL) to the graph neural network (GNN), since GNN can extract spectral information and relations among neighboring pixels.

How to consider the distance between label consistency? Inspired by ADA [162], CPDA [97] presented a circled similarity propagation-based domain adaptation method for the local climate zones (LCZ) classification task, which proposes an adaptation loss based on circled similarity matrix calculated by cosine similarity. The experiments show the effectiveness of the CPDA method on the LCZ42 Dataset [71] using different backbones [4,5,163]. In addition, the Augmented Associative Learning-based (AAL-based) domain adaptation network [164] achieved the hyper-spectral remote sensing image classification by utilizing source classification loss, walking loss and visiting loss.

How to consider the distance between joint distribution consistency? Reference [165] incorporated the label information to align the joint distributions between feature and labels. DTJM [166] is proposed as a domain adaptation method for hyper-spectral images classification through embedding label information of the source domain and designing a discriminative transfer joint method.

However, MLADA [75] mentioned that these methods fail in unsupervised domain adaptation task based on synthetic aperture radar data due to its considerable variations of SAR images in different frequency bands; thus, there are application defects for distance-based methods.

The second part: GAN-based methods.

DANN [27] proposed an architecture including a deep feature extractor, a label predictor and a domain classifier and uses Gradient reversal layer to achieve an end-to-end trainable network. ADDA [28] combined discriminative modeling, untied weight sharing and an adversarial discriminative loss to be trained on unlabeled data from the target domain. Faced with a sea fog detection task based on satellites and meteorological observations, SFGUDA [167] presented a two-stage method which contains an unsupervised domain adaptation module and a seeded region growing module, considering the feasibility of obtaining abundant visible information over the land and the similarity between land fog and sea fog. This work [168] used CycleGAN [114] to capture the transferable features in a same feature subspace. A two-way mapping of source and target domain features is designed and a cycle-consistency is used to minimize the discrepancy of reconstructed features in different domains.

Reference [70] introduced simulated SAR data to synthetic aperture radar automatic target recognition (SAR-ATR) task to solve insufficient labels using adversarial domain adaptation with Wasserstein distance [26] to replace unstable adversarial loss. A neural network trained on an SAR image set of one band is not suitable for the classification of another band images due to the discrepancy of the frequency band. Thus, MLADA [75] proposes a multi-level domain adaptation method based on adversarial learning to solve the domain shift between different band domains, which demonstrate better performance than ADDA (three-feature-level adaptation).

Besides, some works [169] combined GAN-based methods and attention mechanism to improve the models' transferability. For example, MADAN [169] designed a BIN-based feature extractor and introduced a multi-level attention mechanism, which included a feature level attention generated by shallow features and an entropy level attention produced by a deep discriminative feature.

A few work considered characteristics of remote sensing data and tasks so that the features extracted by convolution networks are further transformed to obtain a better performance. For example, Reference [31] converted deep features extracted by a CNN structure into a 2D feature curves and reduces the discrepancy between two curve domains based on a conditional generative adversarial networks (cGANs) model.

3.2.2. Pixel-Level Adversarial Training Methods

Different from Generative training methods migrating discrepancy of appearances and feature-based Adversarial training methods adapted in semantic differences, pixel-level domain adaptation mainly effects on output space to narrow the gap between the prediction of each pixel from the source and target domain. The direct method to solve the above-mentioned problems is designing a domain discriminator for output space to judge where the output is from source or target domain, such as AdaptSegNet [29].

Reference [170] used a fully-convolutional segmentation network as the generator to extract semantic features, and utilizes a discriminator structure to distinguish the prediction results from the target domain based on a binary cross-entropy loss. Besides, papers [32,119] used a similar CLAN method [30] to achieve in road detection for remote sensing data, which defines the category-level feature after element-wise addition and multiplication as global features and decides how well a feature is category-level aligned between source and target by co-training practice.

3.2.3. Hierarchical Adversarial Training Methods

Based on the mining of more effective information, many papers are devoted to realizing multi-level feature adversarial training, including image, region, pixel or semantic level. For example, TA^3N [125] utilizes image-level and region-level domain discriminative features to migrate domain discrepancy and adds transferable attention mechanism in order to focus more on salient objects and less on background. Reference [171] proposed an adversarial adaptive detection network and achieved hierarchical feature adaptation, including image level and semantic level. For the final prediction matrix of the detector, this work combines the Faster-RCNN [6] detector prediction and three Context Modules to enhance feature extraction.

In addition, many works construct a conjunction between distance-based and GAN-based training to achieve adaptation. CsDA [118] embeds GcGAN into a co-training adversarial learning network for land cover mapping using very-high-resolution (VHR) optical aerial images to emphasize the importance of aligning of category-level consistency and global domain consistency.

In recent years, more and more papers have considered different levels fusion from the feature level and pixel level, such as CyCADA [150], AdaptSegNet [29], Advent [172] and so on. Besides, following the transfer of the segmentation output level, there are also many methods for the detection task to use multi-level adversarial training strategy to achieve unsupervised domain adaptation. For example, reference [101] uses prediction and feature alignment to achieve vehicle detection avoiding performance deficits caused by only feature-level migration. CaDA [33] not only considers a joint local and global feature adversarial adaptation on feature and output space, but entropy minimization is also restored in the coastal land cover mapping task.

3.3. Self-Training Methods

With the rise of self-supervised methods, for the self-training method in UDA, there are two main issues worthy of attention, one is how to get better and more effective pseudo labels from models trained on the source domain, and the other is how to use pseudo labels for self training of the target domain.

At present, the acquisition of pseudo labels in unsupervised domain adaptation is mainly following previous works, such as CBST [35] and IAST [36], but most of these methods directly use the idea of pseudo label generation when applied in remote sensing. LoveDA [79] demonstrates the effectiveness of self training methods [35,36,173] with better performance compared other adversarial training methods [23,30] for rural to urban transfer, and vice versa. In addition, it is worthy to mentioned that LoveCS [123] proposed a multi-scale pseudo labeling method to tackle the scale dilemma problem based on CBST and improve the accuracy of segmentation on the LoveDA Dataset.

As for how to utilize the pseudo labels to perform precisely in the target domain, most works train or fine-tune the model obtained from the source domain and some works, such as [121,122], design an easy-hard mechanism to place extra emphasis on the training model.

3.4. Hybrid Training Methods

In order to pursue better model performance to solve the label dilemma and domain differences, most of the current UDA methods in the field of remote sensing images are in the form of a mixture of multiple methods defined as Hybrid training (HT) methods, that is, they are not limited to using only Generative training (GT), Adversarial training (AT) or Self-training (ST) methods. Most of these are multi-stage methods and connect different training parts via input images after style transfer or pseudo labels. Although this type of method can obtain high model accuracy, it is usually accompanied by defects, such as high computational load and low running speed. In practical applications, it is necessary to compromise between accuracy and speed. According to the method of combining two or more ways, we roughly divide the hybrid training method into the following three types and expound in detail the GT-AT, AT-ST and GT-AT-ST methods.

3.4.1. GT-AT Methods

Inspired by FCAN [134], JPRNet [135] consists of a pixel adaptation network (PAN) using CycleGAN [114], in order to transfer one domain image to another domain, and a representation adaptation network (RAN), which contains a similar FCNNet [134], a Segnet and a Discriminator. Experiments demonstrate that the similar images reconstructed by CycleGAN is helpful to learn good representation. Paper [174] presented a novel approach relying on adversarial training of an appearance adaptation network (AAN) jointly with the classification network, which requires only a single adaptation network from D_S to D_T . Besides, a new regularization term and a new criterion for selecting the optimal parameter values are introduced in this paper. Dcan [175] used adaptive instance normalization to achieve Channel-wise Feature Alignment, because the mean and standard deviation of each channel causes image style discrepancy [176].

ResiDualGAN [38] proposes a two-stage method utilizing a residual inspiration and resize module, considering the scale discrepancy of RS images datasets and a scene of the real-to-real translation, where one stage is to carry out an unsupervised neural transfer to obtain fake images through $X_{S-T} = ResiG_{S-T}(X_S) = Resize_{S-T}(G_{S-T}(X_S) + (X_S))$ and another stage is to train a semantic segmentation model.

Concurrent works [150,177] added two domain-specific task networks named F_T and F_S to ensure their predictions maintained semantic consistency based on leveraging image-to-image translation methods.

3.4.2. AT-ST Methods

Paper [161] used combination of coral correlations and self-training and demonstrated the effectiveness of the proposed method for multi-temporal and hyper-spectral remote sensing images. Similar to paper [178], RoadDA [121,122] proposed an unsupervised inter-domain and intra-domain adaptation method for road detection, which first separates the target domain into easy and hard splits using an entropy-based ranking function, and then decreases the inter-domain or intra-domain gap via an adversarial mechanism. Coincidentally, TriADA [81] is conducted adversarial training on feature and output space and class-aware self training methods to generate pseudo labels for the target domain and retrain classification model for very high resolution (VHR) images. Besides, after image style transfer in the first stage, paper [45] used four images to train a segmentation network in a second stage concluding a source image, a target-stylized source image, a reconstructed source image and a target image.

3.4.3. GT-AT-ST Methods

A few of the methods use all three training modes at the same time due to the high complexity of the model and the instability of training. Paper [37] introduced an unsupervised FasterRCNN-based [6] approach on the miniSAR and FARADSAR dataset, containing pixel domain adaptation on data appearance, multi-feature level adaptation and self training using an iterative pseduo label.

4. Other Concerns of UDA in Remote Sensing

In actual remote sensing applications, there are still many practical problems to achieve accurate ground observation. For example, unknown categories may need to be identified during the observation process, or it could be combined with multiple satellites to achieve comprehensive judgment. In this section, the solutions to the scale difference caused by different resolutions is first introduced specifically and then we introduce the methods dedicated to solving open-set and multi-domain unsupervised domain adaptation problems.

4.1. Scale Divergence Problem

Due to the planar nature of the scene and the nadir viewpoint, in the same satellite, the size or resolution of different objects could not change. However, data from different satellites suffer from non-uniform object size because of different resolution and satellite orbit. It is necessary to alleviate the discrepancy caused by the resolution from various satellites or remote sensing data.

Most methods utilize multi-level feature extraction to minimize the negative effects of the resolution. Reference [171] combined the super-resolution module and Adversarial Adaptive Detection Network (AADN). Reference [179] proposed a dual discriminator network with a Scale Attention Module (SAM). Similarly, this paper chose Deeplab v3+ network [8] with the ASPP module as a backbone to reduce the influence of scale changes. Besides, LoveCS [123] used a multi-scale pseduo labels to narrow the domain gap caused by scale divergence through a dense multi-scale decoder.

4.2. Partial or Open-Set Unsupervised Domain Adaptation

The aforementioned domain adaptation methods are mostly shared the same label space, but in reality, there still exist some unknown categories in the recognition or detection task. Thus, partial unsupervised domain adaptation (PDA) [39,40] and open-set unsupervised domain adaptation (OSDA) [180–182] are also challenging tasks based on remote sensing data. The key of PDA methods is to avoid negative transfer due to the mismatch classes. Massive PDA methods achieve this goal through identifying the source instances which belong to the same class in the target domain. For example, Coordinate partial adversarial domain adaptation (CPADA) [39] transfers relevant samples in the same label class with the help of coordinate loss, while the Selective Adversarial Network (SAN) [40] attempts to avoid negative transfer through eliminating the outlier source classes and maximally matching the data distributions in label space. Under the setting of OSDA, it is difficult to align the same and known categories and separate an unknown class from a known class and source domain data. OSDANet [180] designed a network with a $Q + 1$ dimensional classifier used to separate known and unknown classes based on adversarial training methods. Paper [182] proposed a domain adaptation based on paper [181] in a spherical space instead of the prior Euclidean feature space under open set condition. Besides, this paper combines an adversarial training module and pseduo labels to coverage the ship detection network for two benchmark SAR datasets.

4.3. Multi-Domains Unsupervised Domain Adaptation

Most of the existing methods focus on the circumstances that achieve transfer from a single-source to a single-target domain introduced in Section 3, but it is an ideal situation, while in reality, there are many circumstances faced with more than one source or target

domain due to the need for using multiple detection methods or remote sensing data of different modalities.

Multi-source domain to single-target domain adaptation.

The most straightforward solution of multi-source to single-target domain adaptation is to achieve pairing and narrowing differences between each sub-source and target domain, whether it is a difference in appearance or a difference in features. For the feature alignment, MB-Net [183] was the first to use multi-source domain adaption in the field of remote sensing images and proposed a new multiple domain dataset created from four heterogeneous scene datasets. This paper mainly proposes a multi-branch network, which accepts each source domain and target domain pair, respectively, and narrows the distance between different domains by the mean feature.

Besides, some methods treat different source domains as an aggregated overall source domain and implement unsupervised domain adaptation based on this. The core of the aggregation methods is how to generate an aggregated unified domain through target-stylized images from different sub-source domains, such as MADAN [43], or learnable fusion for classifiers and prediction results, such as references [44,184]. MCSN [44] pointed out that it is difficult to achieve the same source domain and target domain categories in MDA, so it proposed multi-complementary source-domain adaptation (MCSDA), which is a union of multi-source domains that share the same categories with the target domain. MCSN [44] firstly performed feature alignment on the source and target domains. Afterwards, it aggregated the results of the source domain as pseudo labels to train the model, such as DCTN [185]. Reference [184] adopted the Minmax entropy approach to achieve the effect of narrowing the discrepancy between the source domain and the target domain, which avoided discriminators in adversarial training with the help of gradient reversal layers. Reference [184] introduced a learnable fusion layer to fuse the prediction results and obtain the final classification result in the target.

Much work has also been devoted to unifying different source domain styles into a standardized style between source and target domains. StandardGAN [21] is a generative lightweight multi-source domain training framework and first uses AdaIN [142] to unify the multi-domain data styles from different cities so that the model can be trained and tested on the data after standardization. Unlike StandardGAN, which has a module for image standardization that needs to be designed separately for each source domain, DaugNet [22] is also based on AdaIN but only uses one common encoder, one decoder and one discriminator to achieve multi-source domain adaptation.

In addition, a few papers discuss the factors causing degradation of performance of multi-source domain models and propose solutions. For example, AMDA [186] is a simple multi-source domain adaptation (MDA) framework to solve the Local Climate Zone (LCZ) problem in remote sensing imagery, which utilizes a weighted loss function to solve the imbalanced problem of data distribution of multi-source domains.

Single-source domain to multi-target domain adaptation.

There are different methods to transfer the multi-source domains to a single target domain, caused by a lack of mature and rich identification experience and knowledge for single-source to multi-target domains adaptation. Paper [46] was the first work discussing the issue for multi-target domain adaptation in remote sensing and building a challenge based on four dataset [59,65,66,187]. In order to avoid negative transfer due to dataset discrepancy and no target annotations in mixed-multi-target domains, this paper used the meta learning method [188] with an encoder-decoder architecture to design a sub-target domain loss. Based on paper [46], reference [189] extended the multi-target domain adaptation conception to single-source-mixed-multiple-target domain adaptation and designed a two-stage network including sub-target adaptation and the source to target adaptation.

Besides, reference [190] thought that the best base classifier for different datasets is hard to select, so Multiple Domain Adaptation Fusion (MDAF) and Multiple Base Classifier Fusion (MBCF) are proposed based on a neighborhood consistency using adaptive weighting mechanism.

4.4. Domain Generalization in Remote Sensing

Domain generalization (DG) is a brand new and more challenging task that requires the target domain to be unseen, which means completely inaccessible. The problem of DG is how to learn generalized and precise feature representation from a number of related source domains where training data is available, and to then successfully apply it to an “unseen” target domain [191–193]. However, there is a rare attention to Domain generalization in remote sensing. MMD-DRCN [193] attempts to solve large-scale and cross-regional oil palm tree detection by DG. Reference [193] uses both classification loss and reconstruction loss to extract more representative features from multiple source domains and aligns these latent features by MMD loss, so as to obtain a model with stronger generalization. Besides, reference [194] proposes a novel feature-selection method utilizing the measure based on kernel embedding of conditional distributions to provoke generalization capabilities, which include the feature relevance term and domain invariance measure.

5. Discussion

5.1. Comparisons of Different UDA Training Methods

In order to discuss how to choose a more suitable unsupervised domain adaptation method faced with new unlabeled data, we obtained experimental results based on the ISPRS Potsdam and Vaihingen datasets for six categories segmentation task, as shown in Table 2, based on different training methods, including Generative training, Adversarial training, Self-training, Hybrid training and other methods. We also provide the results with different backbones, only trained on the source domain and directly tested on the target domain, or directly trained and tested on the target domain. We can draw the conclusion that at present, GT and HT methods mostly perform better compared to other methods. We argue that it is possible that the discrepancy of remote sensing data is larger than nature images. Besides, it is also migrated from the Potsdam dataset to the Vaihingen dataset, but the channel composition of the source domain data used is different; one is merged by the Red, Green and Blue (RGB) band, while another is merged by the Near Infrared, Green and Blue (IRGB) band, and the performance of the same method is also different. Moreover, the results of the IRGB adaptation are slightly better than the RGB adaptation, since the domain gap between the near infrared and visible bands is larger than between different visible bands.

As can be observed from the above results, the discrepancy across domains defined as domain shift are diverse so that they are addressed by different training methods. For instance, the Adversarial training (AT) method mostly resolves the discrepancy in category and feature characteristics caused by the semantic margin, whereas the Generative training (GT) method primarily resolves the discrepancy in data appearance brought about by various detection regions and bands. Therefore, it is necessary to select an appropriate domain adaptation method for different data, considering many factors, such as algorithm performance, computing resources and inference speed. If the source and target data are selected from the same satellite, they show some discrepancy in color performance or pixel distribution; thus, it is not necessary to choose the generative domain adaptation method, as the proposed model can achieve better performance.

Although many works have achieved very good experimental performance on the ISPRS dataset, there is still a certain gap between the model performances. Moreover, in natural datasets and scenes, the effectiveness of the method is generally proved by bidirectional transfer, that is, two sets of experimental performances from Potsdam to Vaihingen and from Vaihingen to Potsdam are given at the same time. However, many works in the field of remote sensing have ignored this aspect.

Table 2. The experimental results based on the ISPRS Potsdam and Vaihingen datasets for six types of segmentation task using different constitution and resolution bands. Bold text represents the best performance in our statistical experimental results, while underlined text represents the second best performance in a certain class of methods.

Method or Ref			Settings		V9-P5		P5-V9			
					IRRG-RGB		RGB-IRRG		IRGB-IRRG	
			Baseline	Ep	F1 Score	mIou	F1 Score	mIou	F1 Score	mIou
Source Only	[15]		BiSeNet	80	-	-	-	-	0.32	0.17
Source Only	[16]		BiSeNet	-	-	-	0.287	0.167	0.438	0.245
Source Only	[16]		DeepLabv3+	-	-	-	0.449	0.245	0.491	0.253
GT	CycleGAN	[118]	DeepLab	-	0.270	0.215	0.298	0.233	-	-
GT	Method [15]	-	BiSeNet	80	-	-	-	-	0.49	0.30
GT	UST-DG [16]	-	DeepLab v3+	-	-	-	0.509	0.359	0.606	0.416
GT	PRC-GAN [117]	-	DeepLab v3+	45	-	-	<u>0.561</u>	0.407	<u>0.661</u>	<u>0.482</u>
GT	Method [116]	-	SegNet	-	-	-	0.7740	0.6132	-	-
AT	CyCADA	[118]	DeepLab	-	0.433	0.326	0.452	0.363	-	-
AT	Method [31]	-	DeepLab v3+	-	-	0.3870	-	-	-	-
AT	AdaptSegNet	[117]	-	-	-	-	0.401	0.321	0.523	0.352
AT	CLAN	[118]	DeepLab	-	0.487	0.408	0.517	0.426	-	-
AT	CsDA [118]	-	DeepLab	200	<u>0.528</u>	<u>0.423</u>	<u>0.545</u>	<u>0.449</u>	-	-
ST	CBST	[118]	DeepLab	-	0.452	0.374	0.460	0.388	-	-
HT	Method [45]	-	MA-FCN	-	-	-	-	-	-	0.437
HT	DACST [49]	-	VGG-16	-	-	-	-	-	-	0.444
HT	TriADA [81]	-	DeepLab v3	-	-	-	0.656	0.497	<u>0.698</u>	<u>0.551</u>
HT	TriADA-CAST [81]	-	DeepLab v3	-	-	-	0.665	0.514	0.712	0.568
HT	FCAN [174]	-	-	50	-	-	<u>0.669</u>	0.535	-	-
Other	SEANet	[117]	-	-	-	-	0.468	0.278	0.557	0.377
Target Only	[195]		DeepLab v3+	-	0.9212	0.8432	-	-	0.8957	0.8147
Target Only	[195]		DC-Swin	-	0.9325	0.8756	-	-	0.9071	0.8322

5.2. Comparisons of UDA Methods between Natural and Remote Sensing Data

By comparing the unsupervised domain adaptation methods applied to natural images and remote sensing images, for which the statistics are shown in Figure 5, we can find that most of the methods applied in the remote sensing field are inspired by the methods of natural images, and make use of the characteristics and inter-domain differences of remote sensing data to design the method extension and to improve the model transferability. However, similar methods represented in a nature scene applied in remote sensing are usually delayed by about two years, since these methods are sorted by publication date and application task. As a result, compared to the growth of techniques in natural datasets, the development of UDA in remote sensing scene is still relatively modest, and there is still much potential for exploration.

The adversarial training method is the cornerstone underpinning UDA research, while the Self-training method is the rising star, especially in segmentation tasks. The hybrid training method is currently the mainstream development direction of UDA, especially when faced with a large domain shift, and a method based on a single training mode cannot obtain a good performance model. Generative training methods applied in a remote sensing scene are more frequent than in a nature scene due to the larger discrepancy of remote sensing imagery compared with that of natural imagery, where the discrepancy is due to various factors, i.e., detection satellite, band, area and other factors.

The UDA method in the remote sensing field compared with the natural scene needs to pay more attention to the multi-domain problem, not only because the obtained data are different due to the limitation of the detection level and the annotation cannot be obtained immediately, but also because by using multi-modal or multi-source data, a more thorough, accurate identification be reached.

6. Conclusions

Unsupervised domain adaptation based on deep learning has shown excellent advantages in a variety of tasks and applications in the remote sensing scene, which can deal with the dilemma that the model performance will deteriorate when faced with unlabeled data from new areas or detection sensors. From the literature databases, such as the IEEE/SCI/Spring database, according to multiple keywords, including remote sensing, unsupervised domain adaptation, deep learning and other related search terms, more than 200 related works published in journals or conferences after 2012 were retrieved. Based on a detailed reading and analysis of these works, we conducted this review of unsupervised domain adaptation methods for remote sensing scenarios.

In this paper, a comprehensive and fine-grained taxonomy for UDA methods is proposed according to different training modes, namely Generative training (GT), Adversarial training (AT), Self-training (ST) and Hybrid training (HT) methods; these methods can effectively solve the domain gap between the source and target domain. Besides, the definitions and methods of partial, open-set, multi-domain UDA and domain generalization are introduced in this paper in order to attract more academic attention and solve practical remote sensing scene problems. Furthermore, we focus on exploring the causes of the domain shift specific to the remote sensing scene, which is divided into factors related to data acquisition and imaging, tasks and annotations and others. Before selecting various and suitable UDA training approaches, it is vital to assess and analyze the factors and outcomes for the domain shift based on different remote sensing data. For instance, if appearance varies between source and target domain, a Generative training method or a Hybrid training method applying GT at a particular stage could be selected as the training process to obtain a satisfactory model.

Limitations. This paper may have some possible limitations. Regarding the applications in remote sensing selected for this review, there are still many practical applications that have not been introduced in detail in this paper. Moreover, in addition to the rapid development of UDA methods based on Deep Learning (DL), non-DL methods have also made many contributions to this field, such as Transfer Component Analysis [155] and other methods. Besides, in the statistical process of model performance of different methods on the ISPRS dataset, we find that sometimes the same method has a different reproduction accuracy in different works and can even have a large domain gap, which may be due to hyper-parameter settings or test data selection. Even though in this paper we choose to analyze the experimental results with open code as much as possible to ensure the accuracy of the experimental results and inductions, this may still have a small impact.

Deficiencies. Although UDA methods in remote sensing have achieved some success, there are still many key issues that remain unsolved in the current UDA. From the perspective of data, UDA still has certain requirements for the data volume and the correlation of datasets from two domains; otherwise, the performance will be considerably compromised. In addition, because the UDA method relies on the supervised information of the source domain, if there is a large gap between in some categories between two domains, the model may eventually have a weak ability to recognize these categories, or even cause the catastrophic forgetfulness, that is, it cannot recognize them at all.

Future Directions. Based on this review, there are still a lot of topics worth exploring, and we provide a few expectations for future work in this field: (1) Try more self-training methods. At present, self-trained UDA has shown sufficient advantages for natural images, but it has not received much attention in the field of remote sensing. (2) Apply multi-modal data or some low-cost annotated data. Due to the lack of sufficient label information, there is still a large gap between the UDA and full supervision. Fortunately, the rich multi-modal data and some semi-supervised or weakly supervised data in the remote sensing field can help us alleviate this problem. If we can utilize them smartly, the performance of the model can be effectively improved. (3) Try to join continuous learning or test-time learning. For continuous earth observation tasks, the model can be updated automatically over time to better solve the problem of domain gap between observations at different times.

(4) Design the specific UDA method for a type of remote sensing satellite. As mentioned above, the domain shift problem in remote sensing is related to the optical and imaging parameters of the sensor, so the specific satellite situation can be considered in the algorithm design. The explanation of the UDA method can also be explored from this point of view. (5) Establish standardized and comprehensive large-scale datasets. Most existing methods are tested on their own datasets and it is not convenient to compare them with each other. Standardized datasets are more conducive to the development of research. (6) Construct a pre-training model based on a large data scale. A good pre-trained model is very important for UDA. A model with rich prior knowledge and strong generalization can better achieve domain adaptation.

Overall, we believe that these practical issues will receive further attention and development in the future.

Author Contributions: Review of methodologies, M.X. and K.C.; writing—original draft preparation, M.X. and K.C.; writing—review and editing, M.W. and C.Z.; visualization, M.X.; supervision, M.W. and J.G.; project administration, M.W., C.Z. and J.G.; funding acquisition, J.G. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by MoE-CMCC “Artificial Intelligence” Project of Ministry of Education-China Mobile Communications Group with grant number No. MCM20190701.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Liu, P. A survey of remote-sensing big data. *Front. Environ. Sci.* **2015**, *3*, 45. [\[CrossRef\]](#)
2. Zhang, L.; Zhang, L.; Du, B. Deep learning for remote sensing data: A technical tutorial on the state of the art. *IEEE Geosci. Remote Sens. Mag.* **2016**, *4*, 22–40. [\[CrossRef\]](#)
3. Chi, M.; Plaza, A.; Benediktsson, J.A.; Sun, Z.; Shen, J.; Zhu, Y. Big data for remote sensing: Challenges and opportunities. *Proc. IEEE* **2016**, *104*, 2207–2219. [\[CrossRef\]](#)
4. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
5. Xie, S.; Girshick, R.; Dollár, P.; Tu, Z.; He, K. Aggregated residual transformations for deep neural networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 1492–1500.
6. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster r-cnn: Towards real-time object detection with region proposal networks. *Adv. Neural Inf. Process. Syst.* **2015**, *28*, 91–99. [\[CrossRef\]](#)
7. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 5–9 October 2015; Springer: Munich, Germany, 2015; pp. 234–241.
8. Chen, L.C.; Zhu, Y.; Papandreou, G.; Schroff, F.; Adam, H. Encoder-decoder with atrous separable convolution for semantic image segmentation. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 801–818.
9. Wang, M.; Deng, W. Deep Visual Domain Adaptation: A Survey. *Neurocomputing* **2018**, *312*, 135–153. [\[CrossRef\]](#)
10. Wilson, G.; Cook, D.J. A survey of unsupervised deep domain adaptation. *ACM Trans. Intell. Syst. Technol. TIST* **2020**, *11*, 1–46. [\[CrossRef\]](#)
11. Tuia, D.; Persello, C.; Bruzzone, L. Domain adaptation for the classification of remote sensing data: An overview of recent advances. *IEEE Geosci. Remote Sens. Mag.* **2016**, *4*, 41–57. [\[CrossRef\]](#)
12. Kellenberger, B.; Tasar, O.; Bhushan Damodaran, B.; Courty, N.; Tuia, D. Deep Domain Adaptation in Earth Observation. In *Deep Learning for the Earth Sciences: A Comprehensive Approach to Remote Sensing, Climate Science, and Geosciences*; Wiley: Hoboken, NJ, USA, 2021; pp. 90–104. [\[CrossRef\]](#)
13. Tasar, O.; Happy, S.; Tarabalka, Y.; Alliez, P. ColorMapGAN: Unsupervised domain adaptation for semantic segmentation using color mapping generative adversarial networks. *IEEE Trans. Geosci. Remote Sens.* **2020**, *58*, 7178–7193. [\[CrossRef\]](#)
14. Tasar, O.; Happy, S.; Tarabalka, Y.; Alliez, P. Semi2I: Semantically consistent image-to-image translation for domain adaptation of remote sensing data. In Proceedings of the IGARSS 2020—2020 IEEE International Geoscience and Remote Sensing Symposium, Waikoloa, HI, USA, 26 September–2 October 2020; pp. 1837–1840.

15. Benjdira, B.; Bazi, Y.; Koubaa, A.; Ouni, K. Unsupervised domain adaptation using generative adversarial networks for semantic segmentation of aerial images. *Remote Sens.* **2019**, *11*, 1369. [\[CrossRef\]](#)
16. Li, Y.; Shi, T.; Chen, W.; Zhang, Y.; Wang, Z.; Li, H. Unsupervised Style Transfer via Dualgan for Cross-Domain Aerial Image Classification. In Proceedings of the IGARSS 2020—2020 IEEE International Geoscience and Remote Sensing Symposium, Waikoloa, HI, USA, 26 September–2 October 2020; pp. 1385–1388.
17. Fu, H.; Gong, M.; Wang, C.; Batmanghelich, K.; Zhang, K.; Tao, D. Geometry-consistent generative adversarial networks for one-sided unsupervised domain mapping. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 16–20 June 2019; pp. 2427–2436.
18. Zhao, D.; Yuan, B.; Gao, Y.; Qi, X.; Shi, Z. UGCNet: An Unsupervised Semantic Segmentation Network Embedded with Geometry Consistency for Remote-Sensing Images. *IEEE Geosci. Remote Sens. Lett.* **2021**, *19*, 1–5. [\[CrossRef\]](#)
19. Li, Z.; Wang, R.; Pun, M.O.; Wang, Z.; Yu, H. Inverse Domain Adaptation for Remote Sensing Images Using Wasserstein Distance. In Proceedings of the 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, Brussels, Belgium, 11–16 July 2021; pp. 2345–2348.
20. Cai, Y.; Yang, Y.; Zheng, Q.; Shen, Z.; Shang, Y.; Yin, J.; Shi, Z. BiFDANet: Unsupervised Bidirectional Domain Adaptation for Semantic Segmentation of Remote Sensing Images. *Remote Sens.* **2022**, *14*, 190. [\[CrossRef\]](#)
21. Tasar, O.; Tarabalka, Y.; Giros, A.; Alliez, P.; Clerc, S. StandardGAN: Multi-source domain adaptation for semantic segmentation of very high resolution satellite images by data standardization. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, Seattle, WA, USA, 14–19 June 2020; pp. 192–193.
22. Tasar, O.; Giros, A.; Tarabalka, Y.; Alliez, P.; Clerc, S. Dagnet: Unsupervised, multisource, multitarget, and life-long domain adaptation for semantic segmentation of satellite images. *IEEE Trans. Geosci. Remote Sens.* **2020**, *59*, 1067–1081. [\[CrossRef\]](#)
23. Tzeng, E.; Hoffman, J.; Zhang, N.; Saenko, K.; Darrell, T. Deep domain confusion: Maximizing for domain invariance. *arXiv* **2014**, arXiv:1412.3474.
24. Sejdinovic, D.; Sriperumbudur, B.; Gretton, A.; Fukumizu, K. Equivalence of distance-based and RKHS-based statistics in hypothesis testing. *Ann. Stat.* **2013**, *41*, 2263–2291. [\[CrossRef\]](#)
25. Sun, B.; Saenko, K. Deep coral: Correlation alignment for deep domain adaptation. In *Proceedings of the European Conference on Computer Vision*; Springer: Amsterdam, The Netherlands, 2016; pp. 443–450.
26. Arjovsky, M.; Chintala, S.; Bottou, L. Wasserstein generative adversarial networks. In Proceedings of the International Conference on Machine Learning, PMLR, Sydney, NSW, Australia, 6–11 August 2017; pp. 214–223.
27. Ganin, Y.; Lempitsky, V. Unsupervised domain adaptation by backpropagation. In Proceedings of the International Conference on Machine Learning, PMLR, Lille, France, 6–11 July 2015; pp. 1180–1189.
28. Tzeng, E.; Hoffman, J.; Saenko, K.; Darrell, T. Adversarial discriminative domain adaptation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 7167–7176.
29. Tsai, Y.H.; Hung, W.C.; Schuler, S.; Sohn, K.; Yang, M.H.; Chandraker, M. Learning to adapt structured output space for semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–20 June 2018; pp. 7472–7481.
30. Luo, Y.; Zheng, L.; Guan, T.; Yu, J.; Yang, Y. Taking a closer look at domain shift: Category-level adversaries for semantics consistent domain adaptation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 16–20 June 2019; pp. 2507–2516.
31. Liu, W.; Su, F. Unsupervised adversarial domain adaptation network for semantic segmentation. *IEEE Geosci. Remote Sens. Lett.* **2019**, *17*, 1978–1982. [\[CrossRef\]](#)
32. Lu, X.; Zhong, Y. A Noval Global-Local Adversarial Network for Unsupervised Cross-Domain Road Detection. In Proceedings of the 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, IEEE, Brussels, Belgium, 11–16 July 2021; pp. 2775–2778.
33. Chen, J.; Chen, G.; Fang, B.; Wang, J.; Wang, L. Class-Aware Domain Adaptation for Coastal Land Cover Mapping Using Optical Remote Sensing Imagery. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 11800–11813. [\[CrossRef\]](#)
34. Kang, G.; Jiang, L.; Yang, Y.; Hauptmann, A.G. Contrastive adaptation network for unsupervised domain adaptation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 16–20 June 2019; pp. 4893–4902.
35. Zou, Y.; Yu, Z.; Kumar, B.; Wang, J. Unsupervised domain adaptation for semantic segmentation via class-balanced self-training. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 289–305.
36. Mei, K.; Zhu, C.; Zou, J.; Zhang, S. Instance adaptive self-training for unsupervised domain adaptation. In *Proceedings of the European Conference on Computer Vision*; Springer: Cham, Switzerland, 2020; pp. 415–430.
37. Shi, Y.; Du, L.; Guo, Y. Unsupervised Domain Adaptation for SAR Target Detection. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 6372–6385. [\[CrossRef\]](#)
38. Zhao, Y.; Gao, H.; Guo, P.; Sun, Z. ResiDualGAN: Resize-Residual DualGAN for Cross-Domain Remote Sensing Images Semantic Segmentation. *arXiv* **2022**, arXiv:2201.11523.
39. Hu, J.; Tuo, H.; Wang, C.; Zhong, H.; Pan, H.; Jing, Z. Unsupervised satellite image classification based on partial transfer learning. *Aerosp. Syst.* **2020**, *3*, 21–28. [\[CrossRef\]](#)

40. Cao, Z.; Long, M.; Wang, J.; Jordan, M.I. Partial transfer learning with selective adversarial networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 2724–2732.
41. Panareda Busto, P.; Gall, J. Open set domain adaptation. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 754–763.
42. Saito, K.; Yamamoto, S.; Ushiku, Y.; Harada, T. Open set domain adaptation by backpropagation. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 153–168.
43. Zhao, S.; Li, B.; Yue, X.; Gu, Y.; Xu, P.; Hu, R.; Chai, H.; Keutzer, K. Multi-source domain adaptation for semantic segmentation. *Adv. Neural Inf. Process. Syst.* **2019**, *32*.
44. Lu, X.; Gong, T.; Zheng, X. Multisource compensation network for remote sensing cross-domain scene classification. *IEEE Trans. Geosci. Remote Sens.* **2019**, *58*, 2504–2515. [\[CrossRef\]](#)
45. Ji, S.; Wang, D.; Luo, M. Generative adversarial network-based full-space domain adaptation for land cover classification from multiple-source remote sensing images. *IEEE Trans. Geosci. Remote Sens.* **2020**, *59*, 3816–3828. [\[CrossRef\]](#)
46. Zheng, J.; Wu, W.; Fu, H.; Li, W.; Dong, R.; Zhang, L.; Yuan, S. Unsupervised mixed multi-target domain adaptation for remote sensing images classification. In Proceedings of the IGARSS 2020—2020 IEEE International Geoscience and Remote Sensing Symposium, Waikoloa, HI, USA, 26 September–2 October 2020; pp. 1381–1384.
47. Li, Y.; Shi, T.; Zhang, Y.; Chen, W.; Wang, Z.; Li, H. Learning deep semantic segmentation network under multiple weakly-supervised constraints for cross-domain remote sensing image semantic segmentation. *ISPRS J. Photogramm. Remote Sens.* **2021**, *175*, 20–33. [\[CrossRef\]](#)
48. Iqbal, J.; Ali, M. Weakly-supervised domain adaptation for built-up region segmentation in aerial and satellite imagery. *ISPRS J. Photogramm. Remote Sens.* **2020**, *167*, 263–275. [\[CrossRef\]](#)
49. Luo, M.; Ji, S. Cross-spatiotemporal land-cover classification from VHR remote sensing images with deep learning based domain adaptation. *ISPRS J. Photogramm. Remote Sens.* **2022**, *191*, 105–128. [\[CrossRef\]](#)
50. Nyborg, J.; Pelletier, C.; Lefèvre, S.; Assent, I. TimeMatch: Unsupervised cross-region adaptation by temporal shift estimation. *ISPRS J. Photogramm. Remote Sens.* **2022**, *188*, 301–313. [\[CrossRef\]](#)
51. Toldo, M.; Maracani, A.; Michieli, U.; Zanuttigh, P. Unsupervised domain adaptation in semantic segmentation: A review. *Technologies* **2020**, *8*, 35. [\[CrossRef\]](#)
52. Csurka, G. Domain adaptation for visual applications: A comprehensive survey. *arXiv* **2017**, arXiv:1702.05374.
53. Csurka, G.; Volpi, R.; Chidlovskii, B. Unsupervised Domain Adaptation for Semantic Image Segmentation: A Comprehensive Survey. *arXiv* **2021**, arXiv:2112.03241.
54. Qin, R.; Liu, T. A Review of Landcover Classification with Very-High Resolution Remotely Sensed Optical Images—Analysis Unit, Model Scalability and Transferability. *Remote Sens.* **2022**, *14*, 646. [\[CrossRef\]](#)
55. Nagananda, N.; Taufique, A.M.N.; Madappa, R.; Jahan, C.S.; Minnehan, B.; Rovito, T.; Savakis, A. Benchmarking Domain Adaptation Methods on Aerial Datasets. *Sensors* **2021**, *21*, 8070. [\[CrossRef\]](#) [\[PubMed\]](#)
56. Zhang, Y.; Deng, B.; Tang, H.; Zhang, L.; Jia, K. Unsupervised multi-class domain adaptation: Theory, algorithms, and practice. In *IEEE Trans. Pattern Anal. Mach. Intell.* **2022**, *44*, 2775–2792. [\[CrossRef\]](#) [\[PubMed\]](#)
57. Gong, J.; Zhang, M.; Hu, X.; Zhang, Z.; Li, Y.; Jiang, L. The design of deep learning framework and model for intelligent remote sensing. *Acta Geod. Cartogr. Sin.* **2022**, *51*, 475–487.
58. Wang, Y.; Zhu, X.X. The sarptical dataset for joint analysis of sar and optical image in dense urban area. In Proceedings of the IGARSS 2018—2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia, Spain, 22–27 July 2018; pp. 6840–6843.
59. Cheng, G.; Han, J.; Lu, X. Remote sensing image scene classification: Benchmark and state of the art. *Proc. IEEE* **2017**, *105*, 1865–1883. [\[CrossRef\]](#)
60. Neumann, M.; Pinto, A.S.; Zhai, X.; Houlsby, N. In-domain representation learning for remote sensing. *arXiv* **2019**, arXiv:1911.06721.
61. Risojević, V.; Stojnić, V. The role of pre-training in high-resolution remote sensing scene classification. *arXiv* **2021**, arXiv:2111.03690.
62. Chan, L.; Hosseini, M.S.; Plataniotis, K.N. A comprehensive analysis of weakly-supervised semantic segmentation in different image domains. *Int. J. Comput. Vis.* **2021**, *129*, 361–384. [\[CrossRef\]](#)
63. Mañas, O.; Lacoste, A.; Giro-i Nieto, X.; Vazquez, D.; Rodriguez, P. Seasonal contrast: Unsupervised pre-training from uncured remote sensing data. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Online, 11–17 October 2021; pp. 9414–9423.
64. Zhou, W.; Newsam, S.; Li, C.; Shao, Z. PatternNet: A benchmark dataset for performance evaluation of remote sensing image retrieval. *ISPRS J. Photogramm. Remote Sens.* **2018**, *145*, 197–209. [\[CrossRef\]](#)
65. Xia, G.S.; Hu, J.; Hu, F.; Shi, B.; Bai, X.; Zhong, Y.; Zhang, L.; Lu, X. AID: A benchmark data set for performance evaluation of aerial scene classification. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 3965–3981. [\[CrossRef\]](#)
66. Yang, Y.; Newsam, S. Bag-of-visual-words and spatial extensions for land-use classification. In Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems, New York, NY, USA, 2–5 November 2010; pp. 270–279.
67. Helber, P.; Bischke, B.; Dengel, A.; Borth, D. Eurosat: A novel dataset and deep learning benchmark for land use and land cover classification. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2019**, *12*, 2217–2226. [\[CrossRef\]](#)

68. Helber, P.; Bischke, B.; Dengel, A.; Borth, D. Introducing eurosat: A novel dataset and deep learning benchmark for land use and land cover classification. In Proceedings of the IGARSS 2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia, Spain, 22–27 July 2018; pp. 204–207.
69. Diemunsch, J.R.; Wissinger, J. Moving and stationary target acquisition and recognition (MSTAR) model-based automatic target recognition: Search technology for a robust ATR. In *Algorithms for Synthetic Aperture Radar Imagery V*; International Society for Optics and Photonics: Bellingham, WA, USA, 1998; Volume 3370, pp. 481–492.
70. Wang, K.; Zhang, G.; Leung, H. SAR target recognition based on cross-domain and cross-task transfer learning. *IEEE Access* **2019**, *7*, 153391–153399. [[CrossRef](#)]
71. Zhu, X.X.; Hu, J.; Qiu, C.; Shi, Y.; Kang, J.; Mou, L.; Bagheri, H.; Häberle, M.; Hua, Y.; Huang, R.; et al. So2Sat LCZ42: A benchmark dataset for global local climate zones classification. *arXiv* **2019**, arXiv:1912.12171.
72. Xia, G.S.; Bai, X.; Ding, J.; Zhu, Z.; Belongie, S.; Luo, J.; Datcu, M.; Pelillo, M.; Zhang, L. DOTA: A large-scale dataset for object detection in aerial images. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 3974–3983.
73. Xu, T.; Sun, X.; Diao, W.; Zhao, L.; Fu, K.; Wang, H. FADA: Feature Aligned Domain Adaptive Object Detection in Remote Sensing Imagery. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 1–16. [[CrossRef](#)]
74. Li, K.; Wan, G.; Cheng, G.; Meng, L.; Han, J. Object detection in optical remote sensing images: A survey and a new benchmark. *ISPRS J. Photogramm. Remote Sens.* **2020**, *159*, 296–307. [[CrossRef](#)]
75. Zhang, W.; Zhu, Y.; Fu, Q. Adversarial deep domain adaptation for multi-band SAR images classification. *IEEE Access* **2019**, *7*, 78571–78583. [[CrossRef](#)]
76. FaradSAR Dataset. Available online: <https://www.sandia.gov/radar/complex-data/> (accessed on 24 July 2022).
77. miniSAR Dataset. Available online: <https://www.sandia.gov/radar/complex-data/index.html> (accessed on 24 July 2022).
78. Azimi, S.M.; Henry, C.; Sommer, L.; Schumann, A.; Vig, E. Skyscapes fine-grained semantic understanding of aerial scenes. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Korea, 27 October–2 November 2019; pp. 7393–7403.
79. Wang, J.; Zheng, Z.; Ma, A.; Lu, X.; Zhong, Y. LoveDA: A Remote Sensing Land-Cover Dataset for Domain Adaptive Semantic Segmentation. *arXiv* **2021**, arXiv:2110.08733.
80. Rottensteiner, F.; Sohn, G.; Jung, J.; Gerke, M.; Baillard, C.; Benitez, S.; Breikopf, U. The ISPRS benchmark on urban object classification and 3D building reconstruction. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2012**, *1*, 293–298. [[CrossRef](#)]
81. Yan, L.; Fan, B.; Liu, H.; Huo, C.; Xiang, S.; Pan, C. Triplet adversarial domain adaptation for pixel-level classification of VHR remote sensing images. *IEEE Trans. Geosci. Remote Sens.* **2019**, *58*, 3558–3573. [[CrossRef](#)]
82. Mnih, V. *Machine Learning for Aerial Image Labeling*; University of Toronto: Toronto, ON, Canada, 2013.
83. Shermeyer, J.; Hogan, D.; Brown, J.; Van Etten, A.; Weir, N.; Pacifici, F.; Hansch, R.; Bastidas, A.; Soenen, S.; Bacastow, T.; et al. SpaceNet 6: Multi-sensor all weather mapping dataset. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, Seattle, WA, USA, 16–18 June 2020; pp. 196–197.
84. Demir, I.; Koperski, K.; Lindenbaum, D.; Pang, G.; Huang, J.; Basu, S.; Hughes, F.; Tuia, D.; Raskar, R. Deepglobe 2018: A challenge to parse the earth through satellite images. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Salt Lake City, UT, USA, 18–22 June 2018; pp. 172–181.
85. Zhu, Q.; Zhang, Y.; Wang, L.; Zhong, Y.; Guan, Q.; Lu, X.; Zhang, L.; Li, D. A Global Context-aware and Batch-independent Network for road extraction from VHR satellite imagery. *ISPRS J. Photogramm. Remote Sens.* **2021**, *175*, 353–365. [[CrossRef](#)]
86. Baumgardner, M.F.; Biehl, L.L.; Landgrebe, D.A. 220 band aviris hyperspectral image data set: 12 June 1992 indian pine test site 3. *Purdue Univ. Res. Repos.* **2015**, *10*, R7RX991C.
87. Salinas Dataset. Available online: http://www.ehu.es/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes#Salinas (accessed on 9 February 2017).
88. Botswana Dataset. Available online: http://www.ehu.es/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes#Botswana (accessed on 10 February 2010).
89. Kennedy Space Center Dataset. Available online: http://www.ehu.es/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes#Pavia_Centre_and_University (accessed on 29 November 2016).
90. Yang, H.L.; Crawford, M.M. Domain adaptation with preservation of manifold geometry for hyperspectral image classification. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *9*, 543–555. [[CrossRef](#)]
91. Washington DC Mall Dataset. Available online: <https://engineering.purdue.edu/~biehl/MultiSpec/hyperspectral.html> (accessed on 24 July 2022).
92. Sun, Z.; Wang, C.; Wang, H.; Li, J. Learn multiple-kernel SVMs for domain adaptation in hyperspectral data. *IEEE Geosci. Remote Sens. Lett.* **2013**, *10*, 1224–1228.
93. Lin, D.; Xu, G.; Wang, X.; Wang, Y.; Sun, X.; Fu, K. A remote sensing image dataset for cloud removal. *arXiv* **2019**, arXiv:1901.00600.
94. Schmitt, M.; Hughes, L.H.; Qiu, C.; Zhu, X.X. SEN12MS—A Curated Dataset of Georeferenced Multi-Spectral Sentinel-1/2 Imagery for Deep Learning and Data Fusion. *arXiv* **2019**, arXiv:1906.07789.
95. Meraner, A.; Ebel, P.; Zhu, X.X.; Schmitt, M. Cloud removal in Sentinel-2 imagery using a deep residual neural network and SAR-optical data fusion. *ISPRS J. Photogramm. Remote Sens.* **2020**, *166*, 333–346. [[CrossRef](#)] [[PubMed](#)]

96. Liu, S.; Shi, Q. Local climate zone mapping as remote sensing scene classification using deep learning: A case study of metropolitan China. *ISPRS J. Photogramm. Remote Sens.* **2020**, *164*, 229–242. [\[CrossRef\]](#)
97. Zhao, N.; Zhong, Y.; Ma, A. Mapping Local Climate Zones with Circled Similarity Propagation Based Domain Adaptation. In Proceedings of the IGARSS 2020—2020 IEEE International Geoscience and Remote Sensing Symposium, IEEE, Waikoloa, HI, USA, 26 September–2 October 2020; pp. 1377–1380.
98. Xu, Y.; Ma, F.; Meng, D.; Ren, C.; Leung, Y. A co-training approach to the classification of local climate zones with multi-source data. In Proceedings of the 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Fort Worth, TX, USA, 23–28 July 2017; pp. 1209–1212.
99. Li, J.; Xu, Z.; Fu, L.; Zhou, X.; Yu, H. Domain adaptation from daytime to nighttime: A situation-sensitive vehicle detection and traffic flow parameter estimation framework. *Transp. Res. Part Emerg. Technol.* **2021**, *124*, 102946. [\[CrossRef\]](#)
100. Koga, Y.; Miyazaki, H.; Shibasaki, R. A method for vehicle detection in high-resolution satellite images that uses a region-based object detector and unsupervised domain adaptation. *Remote Sens.* **2020**, *12*, 575. [\[CrossRef\]](#)
101. Koga, Y.; Miyazaki, H.; Shibasaki, R. Adapting Vehicle Detector to Target Domain by Adversarial Prediction Alignment. In Proceedings of the 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, Brussels, Belgium, 11–16 July 2021; pp. 2341–2344.
102. Zhang, R.; Newsam, S.; Shao, Z.; Huang, X.; Wang, J.; Li, D. Multi-scale adversarial network for vehicle detection in UAV imagery. *ISPRS J. Photogramm. Remote Sens.* **2021**, *180*, 283–295. [\[CrossRef\]](#)
103. Wu, W.; Zheng, J.; Fu, H.; Li, W.; Yu, L. Cross-regional oil palm tree detection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, Seattle, WA, USA, 14–19 June 2020; pp. 56–57.
104. Wu, W.; Zheng, J.; Li, W.; Fu, H.; Yuan, S.; Yu, L. Domain adversarial neural network-based oil palm detection using high-resolution satellite images. *Proc. Autom. Target Recognit XXX SPIE* **2020**, *11394*, 29–37.
105. Zheng, J.; Wu, W.; Zhao, Y.; Fu, H. Transresnet: Transferable Resnet For Domain Adaptation. In Proceedings of the 2021 IEEE International Conference on Image Processing (ICIP), Anchorage, AK, USA, 19–22 September 2021; pp. 764–768.
106. Zheng, J.; Fu, H.; Li, W.; Wu, W.; Yu, L.; Yuan, S.; Tao, W.Y.W.; Pang, T.K.; Kanniah, K.D. Growing status observation for oil palm trees using Unmanned Aerial Vehicle (UAV) images. *ISPRS J. Photogramm. Remote Sens.* **2021**, *173*, 95–121. [\[CrossRef\]](#)
107. Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Fu, C.Y.; Berg, A.C. Ssd: Single shot multibox detector. In *Proceedings of the European Conference on Computer Vision*; Springer: Amsterdam, The Netherlands, 2016; pp. 21–37.
108. Chen, H.; Wu, C.; Du, B.; Zhang, L. DSDANet: Deep Siamese domain adaptation convolutional neural network for cross-domain change detection. *arXiv* **2020**, arXiv:2006.09225.
109. Zhang, C.; Feng, Y.; Hu, L.; Tapete, D.; Pan, L.; Liang, Z.; Cigna, F.; Yue, P. A domain adaptation neural network for change detection with heterogeneous optical and SAR remote sensing images. *Int. J. Appl. Earth Obs. Geoinf.* **2022**, *109*, 102769. [\[CrossRef\]](#)
110. Vega, P.J.S. Deep Learning-Based Domain Adaptation for Change Detection in Tropical Forests. Ph.D. Thesis, PUC-Rio, Rio de Janeiro, Brazil, 2021.
111. Soto, P.; Costa, G.; Feitosa, R.; Happ, P.; Ortega, M.; Noa, J.; Almeida, C.; Heipke, C. Domain adaptation with cyclegan for change detection in the Amazon forest. *ISPRS Arch.* **2020**, *43*, 1635–1643. [\[CrossRef\]](#)
112. Soto, P.J.; Costa, G.A.; Feitosa, R.Q.; Ortega, M.X.; Bermudez, J.D.; Turnes, J.N. Domain-Adversarial Neural Networks for Deforestation Detection in Tropical Forests. *IEEE Geosci. Remote Sens. Lett.* **2022**, *19*, 1–5. [\[CrossRef\]](#)
113. Vega, P.J.S.; da Costa, G.A.O.P.; Feitosa, R.Q.; Adarme, M.X.O.; de Almeida, C.A.; Heipke, C.; Rottensteiner, F. An unsupervised domain adaptation approach for change detection and its application to deforestation mapping in tropical biomes. *ISPRS J. Photogramm. Remote Sens.* **2021**, *181*, 113–128. [\[CrossRef\]](#)
114. Zhu, J.Y.; Park, T.; Isola, P.; Efros, A.A. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017; pp. 2223–2232.
115. Zhang, Y.; Deng, M.; He, F.; Guo, Y.; Sun, G.; Chen, J. FODA: Building change detection in high-resolution remote sensing images based on feature–output space dual-alignment. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 8125–8134. [\[CrossRef\]](#)
116. Schenkel, F.; Middelman, W. Domain adaptation for semantic segmentation of aerial imagery using cycle-consistent adversarial networks. In Proceedings of the 2020 IEEE International Geoscience and Remote Sensing Symposium IGARSS, Waikoloa, HI, USA, 26 September–2 October 2020; pp. 1448–1451.
117. Shi, T.; Li, Y.; Zhang, Y. Rotation Consistency-Preserved Generative Adversarial Networks for Cross-Domain Aerial Image Semantic Segmentation. In Proceedings of the 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, Brussels, Belgium, 11–16 July 2021; pp. 8668–8671.
118. Fang, B.; Kou, R.; Pan, L.; Chen, P. Category-sensitive domain adaptation for land cover mapping in aerial scenes. *Remote Sens.* **2019**, *11*, 2631. [\[CrossRef\]](#)
119. Lu, X.; Zhong, Y.; Zheng, Z.; Wang, J. Cross-domain road detection based on global-local adversarial learning framework from very high resolution satellite imagery. *ISPRS J. Photogramm. Remote Sens.* **2021**, *180*, 296–312. [\[CrossRef\]](#)
120. Zhao, D.; Li, J.; Yuan, B.; Shi, Z. V2RNet: An Unsupervised Semantic Segmentation Algorithm for Remote Sensing Images via Cross-Domain Transfer Learning. In Proceedings of the 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, Brussels, Belgium, 11–16 July 2021; pp. 4676–4679.

121. Shen, W.; Wang, Q.; Jiang, H.; Li, S.; Yin, J. Unsupervised Domain Adaptation for Semantic Segmentation via Self-Supervision. In Proceedings of the 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, Brussels, Belgium, 11–16 July 2021; pp. 2747–2750.
122. Zhang, L.; Lan, M.; Zhang, J.; Tao, D. Stageswise unsupervised domain adaptation with adversarial self-training for road segmentation of remote-sensing images. *IEEE Trans. Geosci. Remote Sens.* **2021**, *60*, 1–3. [\[CrossRef\]](#)
123. Wang, J.; Ma, A.; Zhong, Y.; Zheng, Z.; Zhang, L. Cross-sensor domain adaptation for high spatial resolution urban land-cover mapping: From airborne to spaceborne imagery. *Remote Sens. Environ.* **2022**, *277*, 113058. [\[CrossRef\]](#)
124. Shao, Y.; Li, L.; Ren, W.; Gao, C.; Sang, N. Domain adaptation for image dehazing. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 16–18 June 2020; pp. 2808–2817.
125. Yang, J.; Chen, H.; Xu, Y.; Shi, Z.; Luo, R.; Xie, L.; Su, R. Domain adaptation for degraded remote scene classification. In Proceedings of the 2020 16th International Conference on Control, Automation, Robotics and Vision (ICARCV), Shenzhen, China, 13–14 December 2020; pp. 111–117.
126. Mehta, A.; Sinha, H.; Mandal, M.; Narang, P. Domain-aware unsupervised hyperspectral reconstruction for aerial image dehazing. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, Online, 5–9 January 2021; pp. 413–422.
127. Ebel, P.; Schmitt, M.; Zhu, X.X. Cloud removal in unpaired Sentinel-2 imagery using cycle-consistent GAN and SAR-optical data fusion. In Proceedings of the 2020 IEEE International Geoscience and Remote Sensing Symposium IGARSS, Waikoloa, HI, USA, 26 September–2 October 2020; pp. 2065–2068.
128. Bengana, N.; Heikkilä, J. Improving land cover segmentation across satellites using domain adaptation. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *14*, 1399–1410. [\[CrossRef\]](#)
129. Sekrecka, A.; Wierzbicki, D.; Kedzierski, M. Influence of the sun position and platform orientation on the quality of imagery obtained from unmanned aerial vehicles. *Remote Sens.* **2020**, *12*, 1040. [\[CrossRef\]](#)
130. Zhang, Y.; Feng, Z.; Shi, D. The influence of satellite observation direction on remote sensing image. *J. Remote Sens.* **2007**, *11*, 433. [\[CrossRef\]](#)
131. Girshick, R.; Donahue, J.; Darrell, T.; Malik, J. Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 24–27 June 2014; pp. 580–587.
132. Ge, W.; Yu, Y. Borrowing treasures from the wealthy: Deep transfer learning through selective joint fine-tuning. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 1086–1095.
133. Mazza, A.; Sepe, P.; Poggi, G.; Scarpa, G. Cloud Segmentation of Sentinel-2 Images Using Convolutional Neural Network with Domain Adaptation. In Proceedings of the 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, Brussels, Belgium, 11–16 July 2021; pp. 7236–7239.
134. Zhang, Y.; Qiu, Z.; Yao, T.; Liu, D.; Mei, T. Fully convolutional adaptation networks for semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 19–21 June, 2018; pp. 6810–6818.
135. Shi, L.; Wang, Z.; Pan, B.; Shi, Z. An end-to-end network for remote sensing imagery semantic segmentation via joint pixel-and representation-level domain adaptation. *IEEE Geosci. Remote Sens. Lett.* **2020**, *18*, 1896–1900. [\[CrossRef\]](#)
136. Tuia, D.; Munoz-Mari, J.; Gomez-Chova, L.; Malo, J. Graph Matching for Adaptation in Remote Sensing. *IEEE Trans. Geosci. Remote Sens.* **2013**, *51*, 329–341. [\[CrossRef\]](#)
137. Rakwatin, P.; Takeuchi, W.; Yasuoka, Y. Restoration of Aqua MODIS band 6 using histogram matching and local least squares fitting. *IEEE Trans. Geosci. Remote Sens.* **2008**, *47*, 613–627. [\[CrossRef\]](#)
138. Yaras, C.; Huang, B.; Bradbury, K.; Malof, J.M. Randomized Histogram Matching: A Simple Augmentation for Unsupervised Domain Adaptation in Overhead Imagery. *arXiv* **2021**, arXiv:2104.14032.
139. Agarwal, V.; Abidi, B.R.; Koschan, A.; Abidi, M.A. An overview of color constancy algorithms. *J. Pattern Recognit. Res.* **2006**, *1*, 42–54. [\[CrossRef\]](#)
140. Buchsbaum, G. A spatial processor model for object colour perception. *J. Frankl. Inst.* **1980**, *310*, 1–26. [\[CrossRef\]](#)
141. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial nets. *Adv. Neural Inf. Process. Syst.* **2014**, *27*, 2672–2680. [\[CrossRef\]](#)
142. Huang, X.; Belongie, S. Arbitrary style transfer in real-time with adaptive instance normalization. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 1501–1510.
143. Khaleel, A.; Tasar, O.; Charpiat, G.; Tarabalka, Y. Multi-task deep learning for satellite image pansharpening and segmentation. In Proceedings of the 2019 IEEE International Geoscience and Remote Sensing Symposium IGARSS, Yokohama, Japan, 28 July–2 August 2019; pp. 4869–4872.
144. Murez, Z.; Kolouri, S.; Kriegman, D.; Ramamoorthi, R.; Kim, K. Image to image translation for domain adaptation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 4500–4509.
145. Park, T.; Efros, A.A.; Zhang, R.; Zhu, J.Y. Contrastive learning for unpaired image-to-image translation. In Proceedings of the European Conference on Computer Vision, Glasgow, UK, 23–28 August 2020; pp. 319–345.
146. Han, J.; Shoeiby, M.; Petersson, L.; Armin, M.A. Dual Contrastive Learning for Unsupervised Image-to-Image Translation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Virtual, 19–25 June 2021; pp. 746–755.
147. Taigman, Y.; Polyak, A.; Wolf, L. Unsupervised cross-domain image generation. *arXiv* **2016**, arXiv:1611.02200

148. Badrinarayanan, V.; Kendall, A.; Cipolla, R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 2481–2495. [[CrossRef](#)] [[PubMed](#)]
149. Yi, Z.; Zhang, H.; Tan, P.; Gong, M. Dualgan: Unsupervised dual learning for image-to-image translation. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 2849–2857.
150. Hoffman, J.; Tzeng, E.; Park, T.; Zhu, J.Y.; Isola, P.; Saenko, K.; Efros, A.; Darrell, T. Cycada: Cycle-consistent adversarial domain adaptation. In Proceedings of the International Conference on Machine Learning, PMLR, Stockholm Sweden, 10–15 June 2018; pp. 1989–1998.
151. Zhu, C.; Zhao, D.; Qi, J.; Qi, X.; Shi, Z. Cross-Domain Transfer for Ship Instance Segmentation in SAR Images. In Proceedings of the 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, Brussels, Belgium, 11–16 July 2021; pp. 2206–2209.
152. Benaim, S.; Wolf, L. One-sided unsupervised domain mapping. *Adv. Neural Inf. Process. Syst.* **2017**, *30*. [[CrossRef](#)]
153. Sun, B.; Feng, J.; Saenko, K. Return of frustratingly easy domain adaptation. In Proceedings of the AAAI Conference on Artificial Intelligence, Phoenix, AZ, USA, 12–17 February 2016; Volume 30.
154. Gretton, A.; Borgwardt, K.M.; Rasch, M.J.; Schölkopf, B.; Smola, A. A kernel two-sample test. *J. Mach. Learn. Res.* **2012**, *13*, 723–773.
155. Pan, S.J.; Tsang, I.W.; Kwok, J.T.; Yang, Q. Domain adaptation via transfer component analysis. *IEEE Trans. Neural Netw.* **2010**, *22*, 199–210. [[CrossRef](#)]
156. Wang, Z.; Du, B.; Shi, Q.; Tu, W. Domain adaptation with discriminative distribution and manifold embedding for hyperspectral image classification. *IEEE Geosci. Remote Sens. Lett.* **2019**, *16*, 1155–1159. [[CrossRef](#)]
157. Li, Z.; Tang, X.; Li, W.; Wang, C.; Liu, C.; He, J. A two-stage deep domain adaptation method for hyperspectral image classification. *Remote Sens.* **2020**, *12*, 1054. [[CrossRef](#)]
158. Liu, Z.; Ma, L.; Du, Q. Class-wise distribution adaptation for unsupervised classification of hyperspectral remote sensing images. *IEEE Trans. Geosci. Remote Sens.* **2020**, *59*, 508–521. [[CrossRef](#)]
159. Sun, S.; Gu, Y.; Ren, M. Fine-Grained Ship Recognition from the Horizontal View Based on Domain Adaptation. *Sensors* **2022**, *22*, 3243. [[CrossRef](#)] [[PubMed](#)]
160. Zhu, Y.; Zhuang, F.; Wang, J.; Ke, G.; Chen, J.; Bian, J.; Xiong, H.; He, Q. Deep subdomain adaptation network for image classification. *IEEE Trans. Neural Netw. Learn. Syst.* **2020**, *32*, 1713–1722. [[CrossRef](#)] [[PubMed](#)]
161. Wang, W.; Ma, L.; Chen, M.; Du, Q. Joint correlation alignment-based graph neural network for domain adaptation of multitemporal hyperspectral remote sensing images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 3170–3184. [[CrossRef](#)]
162. Haeusser, P.; Frerix, T.; Mordvintsev, A.; Cremers, D. Associative domain adaptation. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 2765–2773.
163. Hu, J.; Shen, L.; Sun, G. Squeeze-and-excitation networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 7132–7141.
164. Chen, M.; Ma, L.; Wang, W.; Du, Q. Augmented associative learning-based domain adaptation for classification of hyperspectral remote sensing images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *13*, 6236–6248. [[CrossRef](#)]
165. Damodaran, B.B.; Kellenberger, B.; Flamary, R.; Tuia, D.; Courty, N. Deepjdot: Deep joint distribution optimal transport for unsupervised domain adaptation. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 447–463.
166. Peng, J.; Sun, W.; Ma, L.; Du, Q. Discriminative transfer joint matching for domain adaptation in hyperspectral image classification. *IEEE Geosci. Remote Sens. Lett.* **2019**, *16*, 972–976. [[CrossRef](#)]
167. Mengqiu, X.; Ming, W.; Jun, G.; Zhang, C.; Yubo, W.; Zhanyu, M. Sea fog detection based on unsupervised domain adaptation. *Chin. J. Aeronaut.* **2021**, *35*, 415–425.
168. Wang, X.; Li, Y.; Cheng, Y. Hyperspectral image classification based on unsupervised heterogeneous domain adaptation cyclegan. *Chin. J. Electron.* **2020**, *29*, 608–614. [[CrossRef](#)]
169. Zheng, J.; Fu, H.; Li, W.; Wu, W.; Zhao, Y.; Dong, R.; Yu, L. Cross-regional oil palm tree counting and detection via a multi-level attention domain adaptation network. *ISPRS J. Photogramm. Remote Sens.* **2020**, *167*, 154–177. [[CrossRef](#)]
170. Deng, X.; Yang, H.L.; Makkar, N.; Lunga, D. Large scale unsupervised domain adaptation of segmentation networks with adversarial learning. In Proceedings of the 2019 IEEE International Geoscience and Remote Sensing Symposium IGARSS, Yokohama, Japan, 28 July–2 August 2019; pp. 4955–4958.
171. Zhang, T.; Li, W.; Ping, F.; Zhe, e. Adaptive Object Detection for Multi-source Remote Sensing Images. *J. Signal Process.* **2020**, *36*, 1407–1414. (In Chinese)
172. Vu, T.H.; Jain, H.; Bucher, M.; Cord, M.; Pérez, P. Advent: Adversarial entropy minimization for domain adaptation in semantic segmentation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 16–20 June 2019; pp. 2517–2526.
173. Lian, Q.; Lv, F.; Duan, L.; Gong, B. Constructing self-motivated pyramid curriculums for cross-domain semantic segmentation: A non-adversarial approach. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Korea, 27 October–2 November 2019; pp. 6758–6767.
174. Wittich, D.; Rottensteiner, F. Appearance based deep domain adaptation for the classification of aerial images. *ISPRS J. Photogramm. Remote Sens.* **2021**, *180*, 82–102. [[CrossRef](#)]

175. Wu, Z.; Han, X.; Lin, Y.L.; Uzunbas, M.G.; Goldstein, T.; Lim, S.N.; Davis, L.S. Dcan: Dual channel-wise alignment networks for unsupervised scene adaptation. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 518–534.
176. Li, Y.; Wang, N.; Shi, J.; Liu, J.; Hou, X. Revisiting batch normalization for practical domain adaptation. *arXiv* **2016**, arXiv:1603.04779.
177. Chen, Y.C.; Lin, Y.Y.; Yang, M.H.; Huang, J.B. Crdoco: Pixel-level domain transfer with cross-domain consistency. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 16–20 June 2019; pp. 1791–1800.
178. Pan, F.; Shin, I.; Rameau, F.; Lee, S.; Kweon, I.S. Unsupervised intra-domain adaptation for semantic segmentation through self-supervision. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 16–18 June 2020; pp. 3764–3773.
179. Deng, X.; Zhu, Y.; Tian, Y.; Newsam, S. Scale Aware Adaptation for Land-Cover Classification in Remote Sensing Imagery. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, Online, 5–9 January 2021; pp. 2160–2169.
180. Zhang, J.; Liu, J.; Shi, L.; Pan, B.; Xu, X. An open set domain adaptation network based on adversarial learning for remote sensing image scene classification. In Proceedings of the 2020 IEEE International Geoscience and Remote Sensing Symposium IGARSS, Waikoloa, HI, USA, 26 September–2 October 2020; pp. 1365–1368.
181. Gu, X.; Sun, J.; Xu, Z. Spherical space domain adaptation with robust pseudo-label loss. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 16–18 June 2020; pp. 9101–9110.
182. Zhao, S.; Zhang, Z.; Zhang, T.; Guo, W.; Luo, Y. Transferable SAR Image Classification Crossing Different Satellites under Open Set Condition. *IEEE Geosci. Remote Sens. Lett.* **2022**, *19*, 1–5. [[CrossRef](#)]
183. Al Rahhal, M.M.; Bazi, Y.; Abdullah, T.; Mekhalfi, M.L.; AlHichri, H.; Zuair, M. Learning a multi-branch neural network from multiple sources for knowledge adaptation in remote sensing imagery. *Remote Sens.* **2018**, *10*, 1890. [[CrossRef](#)]
184. Al Rahhal, M.M.; Bazi, Y.; Al-Hwiti, H.; Alhichri, H.; Alajlan, N. Adversarial learning for knowledge adaptation from multiple remote sensing sources. *IEEE Geosci. Remote Sens. Lett.* **2020**, *18*, 1451–1455. [[CrossRef](#)]
185. Xu, R.; Chen, Z.; Zuo, W.; Yan, J.; Lin, L. Deep cocktail network: Multi-source unsupervised domain adaptation with category shift. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 3964–3973.
186. Elshamli, A.; Taylor, G.W.; Areibi, S. Multisource domain adaptation for remote sensing using deep neural networks. *IEEE Trans. Geosci. Remote Sens.* **2019**, *58*, 3328–3340. [[CrossRef](#)]
187. Xia, G.S.; Yang, W.; Delon, J.; Gousseau, Y.; Sun, H.; Maître, H. Structural high-resolution satellite image indexing. In Proceedings of the ISPRS TC VII Symposium—100 Years ISPRS, Vienna, Austria, 5–7 July 2010; pp. 298–303.
188. Andrychowicz, M.; Denil, M.; Gomez, S.; Hoffman, M.W.; Pfau, D.; Schaul, T.; Shillingford, B.; De Freitas, N. Learning to learn by gradient descent by gradient descent. *Adv. Neural Inf. Process. Syst.* **2016**, *29*. [[CrossRef](#)]
189. Zheng, J.; Wu, W.; Yuan, S.; Zhao, Y.; Li, W.; Zhang, L.; Dong, R.; Fu, H. A Two-Stage Adaptation Network (TSAN) for Remote Sensing Scene Classification in Single-Source-Mixed-Multiple-Target Domain Adaptation (S^2M^2T DA) Scenarios. *IEEE Trans. Geosci. Remote Sens.* **2021**, *60*, 1–13. [[CrossRef](#)]
190. Wei, H.; Ma, L.; Liu, Y.; Du, Q. Combining multiple classifiers for domain adaptation of remote sensing image classification. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 1832–1847. [[CrossRef](#)]
191. Ghifary, M.; Kleijn, W.B.; Zhang, M.; Balduzzi, D. Domain generalization for object recognition with multi-task autoencoders. In Proceedings of the IEEE International Conference on Computer Vision, Boston, MA, USA, 8–10 June 2015; pp. 2551–2559.
192. Li, H.; Pan, S.J.; Wang, S.; Kot, A.C. Domain generalization with adversarial feature learning. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 5400–5409.
193. Zheng, J.; Wu, W.; Yuan, S.; Fu, H.; Li, W.; Yu, L. Multisource-domain generalization-based oil palm tree detection using very-high-resolution (vhr) satellite images. *IEEE Geosci. Remote Sens. Lett.* **2021**, *19*, 1–5. [[CrossRef](#)]
194. Persello, C.; Bruzzone, L. Relevant and invariant feature selection of hyperspectral images for domain generalization. In Proceedings of the 2014 IEEE International Geoscience and Remote Sensing Symposium IGARSS, Quebec City, QC, Canada, 13–18 July 2014; pp. 3562–3565.
195. Wang, L.; Li, R.; Duan, C.; Zhang, C.; Meng, X.; Fang, S. A novel transformer based semantic segmentation scheme for fine-resolution remote sensing images. *IEEE Geosci. Remote Sens. Lett.* **2022**, *19*, 1–5. [[CrossRef](#)]