



Article

Multiscale Change Detection Domain Adaptation Model Based on Illumination–Reflection Decoupling

Rongbo Fan ^{1,2} , Jialin Xie ³, Jianhua Yang ^{1,2,*} , Zenglin Hong ^{1,2}, Yuqi Xu ⁴ and Hong Hou ⁵¹ School of Automation, Northwestern Polytechnical University, Xi'an 710129, China; fanke@mail.nwpu.edu.cn (R.F.)² Shaanxi Provincial Innovation Center for Geology and Intelligent Remote Sensing Application, Xi'an 710129, China³ School of Aeronautics and Astronautics, Fudan University, Shanghai 200433, China⁴ China Association for Science and Technology Service Center for Societies, Beijing 100038, China⁵ School of Marine Science and Technology, Northwestern Polytechnical University, Xi'an 710129, China; houthong@nwpu.edu.cn

* Correspondence: yangjianhua@nwpu.edu.cn

Abstract: In the change detection (CD) task, the substantial variation in feature distributions across different CD datasets significantly limits the reusability of supervised CD models. To alleviate this problem, we propose an illumination–reflection decoupled change detection multi-scale unsupervised domain adaptation model, referred to as IRD-CD-UDA. IRD-CD-UDA maintains its performance on the original dataset (source domain) and improves its performance on unlabeled datasets (target domain) through a novel CD-UDA structure and methodology. IRD-CD-UDA synergizes mid-level global feature marginal distribution domain alignment, classifier layer feature conditional distribution domain alignment, and an easy-to-hard sample selection strategy to increase the generalization performance of CD models on cross-domain datasets. Extensive experiments conducted on the LEVIR, SYSU, and GZ optical remote sensing image datasets demonstrate that the IRD-CD-UDA model effectively mitigates feature distribution discrepancies between source and target CD data, thereby achieving optimal recognition performance on unlabeled target domain datasets.

Keywords: unsupervised domain adaptation; change detection; illumination and reflection decoupling



Citation: Fan, R.; Xie, J.; Yang, J.; Hong, Z.; Xu, Y.; Hou, H. Multiscale Change Detection Domain Adaptation Model Based on Illumination–Reflection Decoupling. *Remote Sens.* **2024**, *16*, 799. <https://doi.org/10.3390/rs16050799>

Academic Editor: Mohammad Awrangjeb

Received: 15 January 2024

Revised: 8 February 2024

Accepted: 14 February 2024

Published: 25 February 2024



Copyright: © 2024 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

Remote sensing image change detection (CD) involves extracting features from multi-temporal homogeneous and heterogeneous images to identify and analyze variations in terrestrial instances within the same geographic area [1]. The ability of deep neural networks to map nonlinear feature spaces is an effective approach for extracting high-dimensional hidden features and minimizing interference. As a result, deep learning models have become increasingly popular in CD tasks [2–5].

However, differences in imaging modes, spatial resolution, spectral irradiance, and noise perturbations among change detection (CD) datasets can lead to covariate shifts [6]. Furthermore, the heterogeneity of multi-temporal data distributions can hinder the generalization capabilities of CD models [7]. Unfortunately, CD models based on supervised deep learning are highly dependent on the feature distribution of the training data. This leads to significant performance degradation when models trained on labeled datasets (referred to as the source domain) are tested on unlabeled datasets (referred to as the target domain). To improve model robustness, it is necessary to continuously integrate annotated data with a variety of feature distributions. However, annotating remote sensing image CD data is a costly and labor-intensive process. Furthermore, the rapid deployment of models on unannotated data with different feature distributions for urgent tasks is impractical.

In the field of CD, methods to improve performance on unlabeled target domain data primarily include fine-tuning strategies based on prior knowledge and unsupervised domain adaptation (UDA). Fine-tuning strategies typically include pre-training, difference mapping, and post-processing techniques [8–10]. These strategies require a substantial amount of annotated data from the source domain to train the model, and do not directly reduce the disparity in feature distribution between different data domains.

For related but not identical source and target domain data, an effective method to improve the recognition performance of CD model on target domain data is to use UDA, called CD-UDA. CD-UDA reduces the distributional differences between two data domains by extracting the shared invariant feature representations between the source and target domains [11]. For the CD model, there is research aimed at reducing the visual difference between the source and target domains and improving the performance on the target domain through an image translation model. The main methods include CycleGAN [12,13], attention-gate based GANs [14], and multiple discriminators [15]. However, image-to-image translation methods may unintentionally alter the image content, resulting in visual inconsistencies between the original and translated images. Furthermore, this problem is exacerbated by the lack of constraints between pairs of bi-temporal CD data.

Therefore, employing domain feature alignment strategies to achieve CD-UDA is an important research direction. However, the implementation of domain feature alignment in CD-UDA faces several challenges, as follows:

1. Lack of a specific UDA framework for CD. Unlike other per-pixel segmentation models, CD models are bi-stream models designed specifically for bi-temporal data. The differential features generated by CD models have difficulty accurately representing bi-temporal data. In addition, the complexity of land cover in CD data leads to poor intra-class consistency in the features generated by CD models, increasing the difficulty of feature alignment;
2. Imbalance in CD samples. The number of samples in unaltered areas significantly exceeds the number of samples in altered areas. This imbalance can cause CD-UDA models to converge to local optima, incorrectly identifying changed areas as unchanged. This in turn leads to the generation of incorrect pseudo-labels for the target domain.

To address the aforementioned problems, we propose a novel illumination–reflection decoupled change detection domain adaptation (IRD-CD-UDA) framework. The IRD-CD-UDA is able to achieve illumination–reflection feature decoupling and multi-scale feature fusion, effectively reducing the global discrepancies in multi-temporal remote sensing images. Tailored to the specifics of CD-UDA, we have developed a domain alignment strategy for mid-layer global feature marginal distributions, a domain alignment strategy for classifier-layer feature conditional distributions, and an easy-to-hard sample selection strategy based on CD map probability thresholds. The novel contributions of this paper can be summarized as follows:

1. A high-performance feature extraction model with illumination–reflection feature decoupling (IRD) and fusion is proposed. The model consists of a module for extracting low-frequency illumination features, a module for extracting high-frequency reflection features, a module for fusing high-frequency and low-frequency features, and a module for decoding difference features. The IRD model can achieve high-performance supervised CD by decoupling the strongly coupled illumination and reflection features, and provides a backbone model for the UDA-CD strategy;
2. A global marginal distribution domain alignment method for middle layer illumination features is proposed. Utilizing the IRD model to extract shared illumination features from multi-temporal data, this method represents bi-temporal data in both source and target domains, promoting global style alignment and stable model convergence between the domains;

3. A conditional distribution domain alignment method for deep features is designed. By minimizing intra-domain disparities and maximizing inter-domain differences, this approach alleviates the issue of fitting features of different classes into the same feature space, due to the poor intra-class consistency of differential features;
4. An easy-to-hard sample selection strategy based on the CD entropy threshold is proposed. Aiming to obtain reliable pseudo-labels with balanced high confidence, this strategy reduces the transfer failure caused by sample imbalance and promotes stable convergence in CD-UDA.

The remainder of this paper is organized as follows. Section 2 details the basic principles and the latest method of domain adaptation; Section 3 introduces the IRD-CD-UDA model, including the algorithm framework and domain adaptation strategies. Section 4 outlines the experiments conducted using the proposed algorithm on three datasets and discusses the experimental results of the study. The paper is concluded in Section 5.

2. Domain Adaptation

Feature alignment-based UDA methods are an important research direction for CD-UDA. Firstly, the domain D in DA is the subject of learning and contains data samples (samples: $x_i \in X$, labels: $y_i \in Y$). The domain with a large number of labeled samples is usually called the source domain $D_s = \{x_i, y_i\}_{i=1}^{n_s}$, and the direction in which the model needs to transfer is the target domain $D_t = \{x_j, y_j\}_{j=1}^{n_t}$. The feature distribution-based transformation method is based on the given distance measure to determine the feature transformation T . Its optimization function is as follows:

$$f^* = \arg \min_{f \in F} \frac{1}{n_s} \sum_{i=1}^{n_s} l(f(x_i), y_i) + \lambda R(T(n_s), T(n_t)) \quad (1)$$

$$s.t. n_s \neq n_t,$$

where n_s is the number of samples of source domain data; $T(\cdot)$ is the feature mapping function; $f(\cdot)$ is a feature extractor; $l(\cdot)$ is the supervised classification loss on the source domain; $R(T(D_s), T(D_t))$ is the transfer regularization term; and λ is the weight for feature distribution alignment. The idea behind DA is as follows: learn a feature transform $T(\cdot)$ to reduce regularization $R(\cdot)$ and improve similarity.

The transfer method based on feature transformation is directly related to the measurement of probability distribution differences [16]. According to the multi-domain feature distribution difference metrics, DA methods based on feature distribution transformation methods mainly include marginal distribution alignment [17] and conditional distribution alignment [18]. The former assumes that the marginal distributions between domains are different ($P_s(x) \neq P_t(x)$), but the conditional probabilities are the same ($P_s(y|x) \approx P_t(y|x)$). It achieves the transfer by reducing the distance of the marginal probability distribution between two domains, i.e., $D(P_s(y|x), P_t(y|x)) \approx D(P_s(x), P_t(x))$. Common metrics are MMD [19], KL dispersion [20], and CORAL [21].

After development, MKMMD [22] introduced the concept of multi-kernel function mapping into MMD. JDA [23] and VDA [24] introduce conditional probability loss to MMD. DDC [25], DAN [16], and JMMD [26] use MMD for deep features generated by deep learning. BDA [27] and DDAN [28] acquire coefficients by computing the A-distance between source and target domains to harmonize conditional/marginal probability fractions. WDAN [29] appends the sample class fractions in the source domain to the MMD. DGR [30] regulates intraclass and interclass loss fractions using Wasserstein distance. CAN [31] optimizes an MMD to explicitly model intra-class domain disparities and inter-class domain discrepancies. In the area of CD, Chen et al. [32] first proposed the use of multi-kernel maximum mean difference (MK-MMD) to manipulate the disparity features generated by the Siamese fully connected network in order to achieve alignment between the source and target domain disparity features. The selection of high-confidence shared features

between multiple domains is key to improving the efficiency and performance of DA. In general, sample selection strategies are combined with measures of feature distribution differences, such as migration strategies based on centroid distance iteration [33]; migration schemes based on centroid memory mechanisms [34]; and migration strategies based on course management [35].

Based on the above introduction, the UDA method based on feature alignment mainly consists of high-confidence sample screening and to-be-aligned feature extraction with domain feature difference metric calculation. Therefore, our work is mainly based on these three parts and is targeted at CD-UDA.

3. Methodology

The proposed multi-scale change detection unsupervised domain adaptation model based on illumination–reflection decoupling (IRD-CD-UDA) is shown in Figure 1. The model implements an end-to-end CD-UDA framework that includes a mid-deep layer domain alignment strategy and a target domain bi-temporal high-confidence pseudo-labeled sample selection strategy.

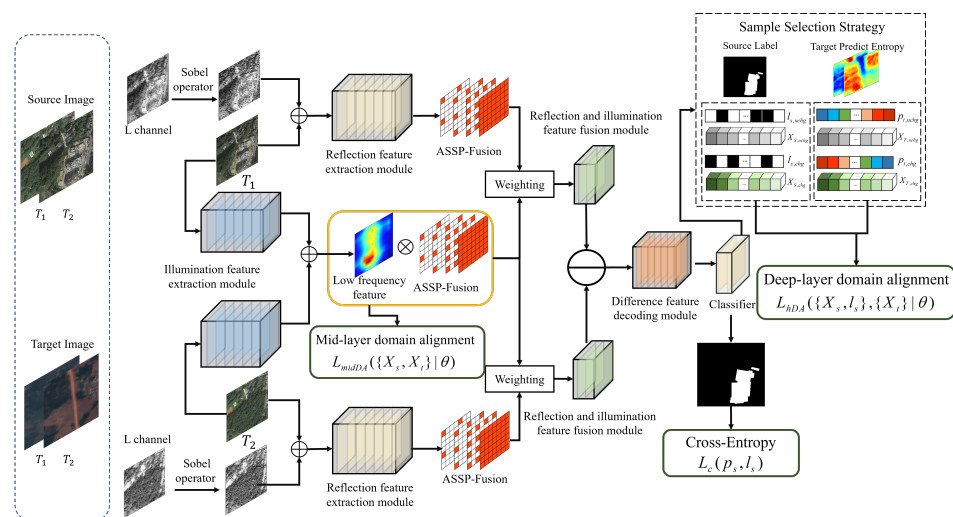


Figure 1. Framework of the IRD-CD-UDA algorithm.

The IRD-CD-UDA is comprised of the following five sections: 1. The low-frequency illumination module extracts the shared global features of the two temporal images (Section 3.1.1) and is also used to align the marginal cross-domain distributions in the middle layer (Section 3.2.1); 2. The high-frequency reflection module extracts the content features of the two temporal data points (Section 3.1.2); 3. The Atrous Apatial Pyramid Pooling (ASSP)-based fusion module [36] is utilized to achieve the multi-scale fusion of illumination and reflection features (Section 3.1.3); 4. Using the difference feature decoding module to map the fused difference feature into a CD probability map (Section 3.1.4), where the classifier layer features are used for deep domain conditional distribution alignment (Section 3.2.2); 5. An easy-to-hard sample selection strategy based on entropy is implemented to achieve CD-UDA transferable layer feature selection and generation (Section 3.2.3).

3.1. Structure of the Illumination–Reflection Decoupled Change Detection Multi-Scale Unsupervised Domain Adaptation Model (IRD-CD-UDA)

For global disturbances affecting multi-temporal remote sensing images, such as illumination, shading, and color distribution differences, they can cause significant distortions in feature distributions at different times within the same region, and lead to the CD model not being able to effectively extract robust difference features [37,38]. Therefore, the CD model with illumination–reflection decoupling is established by relying on the Retinex

theory [39] in image enhancement algorithms. According to the Retinex theory, an image can be described as the product of illumination and reflectance components, where the reflectance component usually represents an intrinsic property of the object. The general formulation of the Retinex theory is as follows:

$$I(x, y) = R(x, y) \times L(x, y), \quad (2)$$

where $I(x, y)$ represents the original image captured by the sensors, $L(x, y)$ represents the illumination component, which reflects the overall illumination information of the image, and $R(x, y)$ is the reflectance component, which is used to characterize the inherent properties of the object. The main factors determining the effectiveness of remote sensing imaging are the reflectivity of ground objects and lighting conditions. Assuming that the satellite is located at the zenith and not taking into account the scattering and absorption effects of the atmosphere, the reflectance is the ratio of the intensity of the reflected radiation received by the sensor from the ground target to the intensity of the solar radiation received by the ground target itself.

Based on the aforementioned idea, considering the problem of coupling between illumination and reflectance features in remote sensing images, a deep network model is employed to achieve image feature decoupling. The schematic diagram of the model, shown in Figure 1, includes a low-frequency illumination feature extraction module, a high-frequency reflectance feature extraction module, an illumination and reflectance feature fusion module, and a difference feature decoding module.

3.1.1. Illumination Feature Extraction Module

The role of the low-frequency illumination feature extraction module consists of three main parts. First, it aims to correct the global illumination features of two remote sensing images by ignoring the discrete high-frequency difference information, so as to effectively mitigate the distribution differences caused by global factors (illumination and atmosphere). Second, the module generates global representation features of the source or target domain to provide domain-discriminative information for CD-UDA, and minimizes cross-domain global differences in subsequent convolutional neural network processing. Finally, the bi-temporal shared features generated by the model in the intermediate layer are used to achieve cross-domain marginal distribution alignment, which improves overall performance.

The structure diagram of the low-frequency illumination feature extraction module is shown in Figure 2. The low-frequency illumination feature extraction module is a two-stream structure with shared parameters. The operational steps are as follows: inputting multi-temporal images $\{X_1, X_2\}$; using convolutional networks to realize feature extraction, which includes convolutional operation with a kernel of seven to realize feature extraction with a large receptive field; finally, the generated deep features are inputted to the spatial attention module [40] to obtain the low-frequency illumination features $X^L \in R^{1 \times \frac{H}{8} \times \frac{W}{8}}$, where H and W are the height and width of the feature, respectively.



Figure 2. Structure of low-frequency illumination feature extraction network.

3.1.2. Reflection Feature Extraction Module

The high-frequency reflection feature extraction module is the backbone module of the IRD-CD model and is composed of residual structures [41]. It aims to improve the CD model's ability to identify changed regions by emphasizing the edges of critical feature targets through high-frequency texture feature weighting operations. As shown in Algorithm 1, the operational steps for the high-frequency reflection feature extraction module are as follows: 1. Convert the multi-temporal data to the Lab color space separately, extract the L channel to obtain $\{X_{L1}, X_{L2}\}$; 2. Convolve $\{X_{L1}, X_{L2}\}$ with two fixed parameter Sobel operators to derive the high-frequency texture features $\{X_{Ln}^{Sx}, X_{Ln}^{Sy}\}, n = 1, 2$ in both X and Y directions; 3. Use the square root operation to compute the texture weighting coefficients $\{X_{L1}^W, X_{L2}^W | X_{Ln}^W = \sqrt{(X_{Ln}^{Sx})^2 + (X_{Ln}^{Sy})^2}, n = 1, 2\}$; 4. Apply the weighting to $\{X_1, X_2\}$, yielding $\{X_1^{WI}, X_2^{WI} | X_n^{WI} = X_n + X_n^W, n = 1, 2\}$; 5. Obtain the high-frequency texture features $\{X_1^H, X_2^H | X_n^H \in R^{128 \times \frac{H}{8} \times \frac{W}{8}}\}$ by using a residue structure.

Algorithm 1 Reflection feature extraction module

Require: multi-temporal images $\{X_1, X_2\}$.

High frequency weighting:

- 1: calculate the L channel $\{X_{L1}, X_{L2}\}$;
- 2: calculate high-frequency texture features $\{X_{Ln}^{Sx}, X_{Ln}^{Sy}\}, n = 1, 2$ by Sobel operators;
- 3: calculate texture weighting factor $\{X_{L1}^W, X_{L2}^W\}$;
- 4: weight high frequency feature $\{X_1^{WI}, X_2^{WI}\}$;

Feature Extraction: obtain the high-frequency texture features $\{X_1^H, X_2^H | X_n^H \in R^{128 \times \frac{H}{8} \times \frac{W}{8}}\}$ by using a residual structure.

Output: $\{X_1^H, X_2^H\}$

3.1.3. Reflection and Illumination Feature Fusion Module

After separately extracting high-frequency and low-frequency features, the feature fusion module is employed to execute the weighted fusion of $\{X_n^L, X_n^H\}, n = 1, 2$. Given the considerable variation in feature target sizes within remote sensing images, the atrous spatial pyramid pooling (ASPP) [36] module is utilized to achieve multi-scale fusion of features, thereby enhancing the completeness of multi-scale feature fusion. The operation steps are as follows: X_n^H and X_n^L are input to the ASSP module to obtain $X_n^{Hassp} = \varphi(\{x_1^H, \dots, x_k^H\})$ and $X_n^{Lassp} = \varphi(\{x_1^L, \dots, x_k^L\})$, respectively; then the fused features $\{X_1^F, X_2^F | X_n^F = conv(X_n^{Hassp} \cdot (1 + X_n^{Lassp})), n = 1, 2\}$, are obtained by multi-scale weighting, splicing, and convolution operations.

3.1.4. Differential Feature Decoding Module

The difference feature decoding module is designed to map fused features $\{X_1^F, X_2^F\}$ into difference features and decode them into CD probabilistic maps. The module uses deconvolution and residual blocks for feature mapping. A skip connection between the high-frequency texture feature extraction module and the differential feature mapping module is incorporated to facilitate cascading between layers with the same subsampling and coding scales. This method utilizes the spatial details in the shallow layers of the network to complement the more abstract and less localized information in the coded data, thus improving the sensitivity to the difference features. After three deconvolution iterations, the change prediction probability map $f_{out} \in R^{B, 2, W, H}$ is finally obtained.

During training, the IRD-CD model utilizes cross-entropy loss L_c to assess the discrepancy between the model's output results and the true distribution outcomes.

$$L_c(p_s, l_s) = -[p_s \cdot \log(l_s) + (1 - p_s) \cdot \log(1 - l_s)], \quad (3)$$

where p_s and l_s are the prediction result of CD and the CD source domain label.

3.2. Strategy of Change Detection Unsupervised Domain Adaptation Model (CD-UDA)

Based on the CD model described in Section 1, the proposed CD-UDA strategy encompasses the following three parts: firstly, a domain alignment strategy targeting the marginal distribution of features in the middle layer; secondly, a domain alignment strategy targeting the conditional distribution of features in the classification layer; thirdly, an easy-to-hard sample selection strategy based on an entropy threshold.

3.2.1. Marginal Distribution Domain Alignment of Illumination Feature

Different domain datasets have different global features such as illumination and atmosphere [37], and the global style features are stable and consistent within a domain. Therefore, the use of the low-frequency illumination feature extraction module (Section 3.1.1) is proposed, in order to extract the global style features in the source and target domains, and realize the marginal distribution domain alignment, which ultimately achieves the purposes of inter-domain global feature alignment and stabilized model convergence.

Although the global illumination feature X^L can serve as one of the features for domain discrimination, there are significant differences in the global features even for the same domain data (same dataset) due to the complex atmospheric illumination. In such cases, enforcing the global features to narrow between different domains will lead X^L to converge to a constant value, which is not physically meaningful. Therefore, the similarity between the global features $\{X_S^L, X_T^L\}$ of the metric source and target domains, rather than the absolute difference, is utilized to align the interlayer features. The style similarity measure is implemented using the Frobenius norm, which computes the low-level features of different images, borrowed from the image style migration model. Notably, there is also prior work in the DA field that leverages the Frobenius norm to achieve domain alignment [21].

Based on the above method, the dimensions of $\{X_S^L, X_T^L\}$ are first converted to $\{X_S^L \in R^{1 \times \frac{H+W}{8}}, X_T^L \in R^{1 \times \frac{H+W}{8}}\}$, respectively. Then, they are normalized to obtain $\widetilde{X}^L = X^L - \overline{X}^L$. The optimization function for the midlayer domain alignment is as follows:

$$\begin{aligned} L_{midDA} &= \frac{1}{4n_t^2} \|C_S - C_T\|_F^2 \\ C_S &= \frac{1}{n_s-1} \left(\left(\widetilde{X}_S^L \right)^T \left(\widetilde{X}_S^L \right) - \frac{1}{n_s} \left(1^T \widetilde{X}_S^L \right)^T \left(1^T \widetilde{X}_S^L \right) \right), \\ C_T &= \frac{1}{n_t-1} \left(\left(\widetilde{X}_T^L \right)^T \left(\widetilde{X}_T^L \right) - \frac{1}{n_t} \left(1^T \widetilde{X}_T^L \right)^T \left(1^T \widetilde{X}_T^L \right) \right) \end{aligned} \quad (4)$$

where n_t is the number of samples in the target domain, and the global difference matrix $A = C_S - C_T, A \in R^{n \times m}$ between the source and target domains is set. The Frobenius norm $\|\cdot\|_F$ is formulated as follows:

$$\|A\|_F = \sqrt{\sum_{i=1}^n \sum_{j=1}^m a_{i,j}^2}, a_{i,j} \in A. \quad (5)$$

3.2.2. Conditional Distribution Domain Alignment of Classification Feature

As deep neural network models deepen, their features evolve from general to specific features. Shallow features enable the characterization of the global style, but they cannot extract crucial feature information for surface instances. The proposed deep feature domain alignment method utilizes the input features $\{X_S^C, X_T^C\}, X_S^C \in R^{32 \times H, W}, X_T^C \in R^{32 \times H, W}$ of the classification layer to achieve conditional distribution alignment based on the pseudo-labels output by the CD classifier.

However, due to the gap between the learned difference features and the large variability of the distance in the Euclidean space between domains, the domain similarity cannot

be appropriately characterized, leading to the model's convergence to a local minimum. To address this issue, the cosine distance is used to measure the difference between domains. Specifically, the feature vector of a point in the source domain image on the transferable feature is denoted as $x_s \in X_s^C$, and the feature vector of a point in the target domain is represented as $x_t \in X_t^C$. The inter-domain difference matrix $D_{st}(x_s^i, x_t^j)$ is

$$D_{st}(x_s^i, x_t^j) = \frac{(x_s^i)(x_t^j)^T}{\|x_s^i\| \times \|x_t^j\|}, i \in (0, n_s), j \in (0, n_t), \quad (6)$$

where n_s and n_t are the number of samples in the source and target domains. Since $P(X_s) \neq P(X_t)$, the difference features between the changed and unchanged regions exhibit significantly different feature distributions. If the difference between the two domains is directly reduced using marginal distribution alignment, features from different classes will be fitted into the same feature space. The specificity of the features between the different classes would be reduced, resulting in lower classification performance.

To address the above issues, we propose a conditional distribution adaptation, to minimize intra-class domain difference and maximize inter-class domain difference. To prevent transfer failure caused by fitting features from different classes to the same feature space, the labels of the source domain are denoted as $l_s = \{l_{s,uchg}, l_{s,chg}\} \in R^{n_s \times 2}$, where $l_{s,uchg}$ and $l_{s,chg}$ represent the changed and unchanged labels of the source domain data, respectively. The CD probability of the target domain obtained by IRD-CD is represented as $p_t = \{p_{t,uchg}, p_{t,chg}\} \in R^{n_t \times 2}$, where $p_{t,uchg}$ and $p_{t,chg}$ denote the probabilities of the changed and unchanged classification results of the target domain data fed into the source domain training model, respectively. The pseudo-label of the target domain is $l_t \in R^{n_t}$. The domain weight matrix $M_{ts} \in R^{(n_s+n_t) \times (n_s+n_t)}$ is as follows:

$$M_{ts} = \begin{bmatrix} p_{t,uchg} l_{s,uchg}^T & p_{t,chg} l_{s,chg}^T \end{bmatrix}. \quad (7)$$

The adaptive weighted difference feature DA loss L_{hDA} is defined as follows:

$$L_{hDA}(\{X_s, l_s\}, \{X_t\} | \theta) = M_{ts} \cdot D(f_\theta(x_s^i), f_\theta(x_t^j)), \quad (8)$$

where $D(\cdot)$ is the Euclidean distance.

3.2.3. Sample Selection Strategy

There is frequently an imbalance between unchanged and changed regions in CD samples, resulting in a higher number of unchanged instances. The sample imbalance causes the CD model to easily converge to a local minimum, which can misidentify the changed regions as unchanged regions and reduce the recall of the classifier. The above problem is particularly prominent in the CD-UDA domain. In addition, in convolution-based CD networks, a model with one data batch provides a large amount of deep feature data, leading to excessive computational intensity. Therefore, as shown in Algorithm 2, an easy-to-hard sample selection strategy is proposed, which can be used to extract effective target domain data by manipulating the entropy threshold of the CD map.

Algorithm 2 Transferable sample selection strategy**Require:** The transferable features from source and target domain $\{f_s, f_t\}$.**Initialization:** Set the entropy threshold T_p , sample number to be selected n_s and n_t .**Transferable sample selection:**

- 1: the samples of $E_t > T_p$ in f_t are selected, and the number is counted as n_{t0} ;
- 2: if $n_{t0} < n_t$ then $T_{p'} = v \cdot T_p$;
- 3: select the features with $E_t > T_{p'}$ in f_t and count their number as n_{t1} . Get the target samples f_t' by concatenating;
- 4: the number of n_{t1} transferable features are randomly selected in f_s , and then the samples are concatenated to obtain f_s' .

Output: Transferable samples $\{f_s', f_t'\}$.

Entropy can be used as a measure of the model's prediction confidence on samples from the target domain. For the CD task, a low entropy value usually indicates that the model is relatively more certain about its predictions, while a high entropy value indicates a high degree of uncertainty in the predictions. The transferable features within the source and target domains are denoted as $f_s \in R^{b \times c \times h \times w}$ and $f_t \in R^{b \times c \times h \times w}$, respectively. Based on the target domain classification probability $\{p_t^{unchg}, p_t^{chg}\}$, the entropy E_t of the target domain is calculated as follows:

$$E_t^{chg} = -p_t^{chg} \log_2(p_t^{chg})$$

$$E_t^{unchg} = -p_t^{unchg} \log_2(p_t^{unchg}) \quad (9)$$

where E_t^{chg} and E_t^{unchg} are the entropy values of the changed and unchanged regions, respectively. Features greater than the entropy threshold T_p are selected as $\{f_t^{chg_i}, f_t^{unchg_i} | E_t^i > T_p, f_t^i \in R^c, i = 1, \dots, n_{t0}\}$ from the target domain samples.

However, the imbalance of samples leads to a decrease in the confidence E_t^{chg} of the samples in the changed region predicted by the CD model, resulting in a large number of samples in the changed region being incorrectly discarded. Therefore, if the number of samples n_{t0} obtained from the first entropy screening is less than n_t (a preset parameter), the probability threshold of the changed region for the second screening is reduced to $v \cdot T_p$, and $\{f_t^{chg_i}, f_t^{unchg_i} | E_t^{unchg_i} > v \cdot T_p, E_t^{chg_i} > T_p, f_t^i \in f_t, i = 1, \dots, n_{t1}\}$ is obtained where the number of samples is n_{t1} . The n_{t1} source domain features are randomly selected from the source domain samples. The source domain for feature alignment is $f_s' = [f_s^{chg_0}, \dots, f_s^{chg_n}, f_s^{unchg_0}, \dots, f_s^{unchg_m}]$, $f_s' \in R^{(n+m)c}$ and the target domain features are $f_t' = [f_t^{chg_0}, \dots, f_t^{chg_a}, f_t^{unchg_0}, \dots, f_t^{unchg_b}]$, $f_t' \in R^{(a+b)c}$ after stitching. To overcome the effect of sample imbalance on domain alignment, the number of selected changed/unchanged features is set to $n = m = a = b$. Figure 3 shows the results of the sample selection.

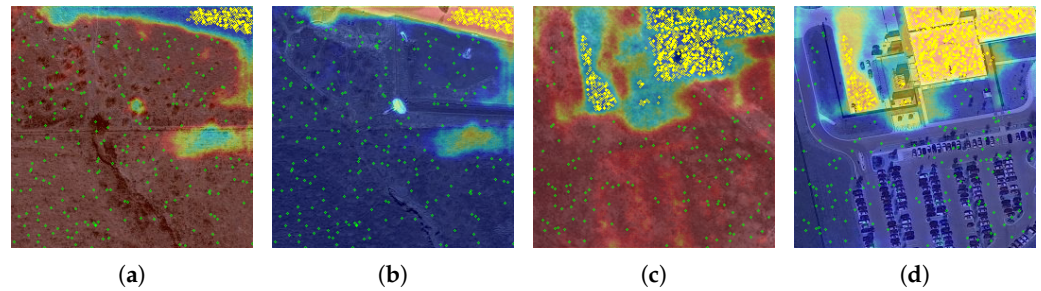


Figure 3. The results of the proposed sample selection strategy; the yellow dots are the changed region sample selection points, the green area is the unchanged region sample selection points. (a,b) are the LEVIR dataset's multi-temporal phase data; (c,d) are the GZ dataset's multi-temporal phase data; (a,c) is the unchanged region heat mask map; and (b,d) is the changed region heat mask map.

In summary, the algorithm flow of IRD-CD-UDA is shown in Algorithm 3. The overall optimization function is as follows:

$$L_{DA} = L_c + L_{midDA} + L_{hDA}. \quad (10)$$

Algorithm 3 Framework of the IRD-CD-UDA algorithm

Require: Multi-temporal data from source domain and target domain: $\{I_S, I_T\} \in R^{w \times h \times c}$.
Initialization: randomly set parameter of IRD-CD Θ follow the standard normal distribution.

Normalization: $\{\hat{I}_S, \hat{I}_T\}$.

Pretraining based on $\hat{I}_S$:

1: Using $\hat{I}_S$ and p_s , Update the Θ by minimizing L_c (3);

Unsupervised domain adaptation of CD:

2: The classification layer transferable feature $\{X_S^C, X_T^C\}$ and mid-layer transferable feature $\{X_S^L, X_T^L\}$ are obtained by $\{\hat{I}_S, \hat{I}_T, \Theta\}$;

3: Use the sample selection strategy from easy to hard based on probability threshold to get $\{f_s^{chg_i}, f_s^{chg_i}, f_t^{unchg_i}, f_s^{unchg_{im}}\}$;

4: The mid-layer feature marginal distribution domain adaptation by (4);

5: Classification layer feature conditional distribution domain adaptation by (8);

6: Update the Θ by (10) until converged;

Output: CD results on the target domain $\hat{I}_T$.

4. Experiment

4.1. Experimental Setting and Datasets

The experiments are set up in the following three parts: first, the experiments on the performance of the IRD-CD model, which were proposed in Section 3.1, to verify the effectiveness of the proposed improvements; second, ablation experiments on the proposed UDA strategy, to verify the effectiveness of the UDA strategy; third, comparing the domain adaptive methods for change detection and adding the CD-UDA strategy to the better performing CD model, in order to realize the generalization of the proposed CD-UDA strategy for validation. An NVIDIA 3090 24 GB graphics processor is utilized as the hardware platform for model training with a batch size of 10. The learning rate is 10^{-3} , and training is performed for five epochs. Table 1 provides an overview of the experimental dataset, where P_c and P_u represent the proportion of samples in the changing and unchanged regions.

Table 1. Overview information of experimental datasets.

Name	Source	Ground Target Composition	$P_c:P_u$
SYSU [42]	Aerial images	Urban construction, suburbs, pre-construction, vegetation, road, ocean	0.20:1
GZ [43]	Google Earth	Water bodies, roads, farmland, bare land, forests, buildings, ships	0.10:1
LEVIR [44]	Google Earth	Buildings, including urban and rural scenes	0.05:1

4.2. Evaluation Metrics

The experimental evaluation metrics [10] include overall accuracy (OA), change area accuracy (CAcc), unchanged area accuracy (UAcc), mean cross-ratio union (mIoU), and F1 score (F1) in the confusion matrix. The imbalanced CD samples make it difficult to evaluate model performance from the OA metric alone, because even a high OA has no application value if a large number of samples in the class to which the model is fitted are at extreme values (local minima). Therefore, the experimental analysis will focus on F1 and mIoU [45].

4.3. Performance Evaluation of Change Detection

Since UDA-CD is an extension and enhancement of CD, the performance of the CD model has a direct impact on the migratability effect of the CD-UDA model. Therefore, the performances of supervised IRD-CD models need to be evaluated. Six methods with stable and advanced performance are selected for comparative experiments, including Deeplab [46], FCSiamD [47], EUNet [48], UCDNet [49], ISNet [50] and proposed IRDNet. These comparison experiments are performed separately for each dataset, and the results are presented in Table 2, with the best results highlighted in bold.

The supervised CD methods for which comparative experiments are conducted are robust and have excellent performance. The results show that their overall performance is relatively close. Among these methods, DeepLab and EUNet (both single-stream models) show suboptimal performance on all three datasets, confirming the idea that the two-stream Siamese model is more suitable for the CD task. Compared to the other concatenated models, the proposed IRD-CD model achieves the best performance on all three datasets, with a clear advantage on the most imbalanced GZ dataset.

Table 2. Performance of six CD models on three datasets.

Data	Method	OA	UAcc	CAcc	F1	mIoU	Parp
GZ	DeepLab	0.967	0.992	0.679	0.876	0.795	26.7M
	FCSiamD	0.980	0.991	0.870	0.933	0.884	2.1M
	EUNet	0.968	0.999	0.627	0.872	0.794	2.3M
	UCDNet	0.964	0.979	0.800	0.884	0.805	4.3M
	ISNet	0.968	0.969	0.951	0.907	0.839	12.3M
	IRDNet	0.985	0.989	0.933	0.909	0.908	6.3M
LEVIR	DeepLab	0.994	0.998	0.733	0.927	0.838	26.7M
	FCSiamD	0.995	0.998	0.789	0.921	0.863	2.1M
	EUNet	0.992	0.997	0.637	0.856	0.775	2.3M
	UCDNet	0.995	0.999	0.774	0.920	0.862	4.3M
	ISNet	0.994	0.996	0.895	0.918	0.859	12.3M
	IRDNet	0.996	0.998	0.847	0.936	0.886	6.3M
SUZYU	DeepLab	0.895	0.953	0.687	0.824	0.732	26.7M
	FCSiamD	0.897	0.942	0.730	0.844	0.741	2.1M
	EUNet	0.864	0.944	0.575	0.781	0.661	2.3M
	UCDNet	0.888	0.918	0.780	0.840	0.734	4.3M
	ISNet	0.893	0.954	0.670	0.832	0.726	12.3M
	IRDNet	0.905	0.962	0.697	0.851	0.751	6.3M

4.4. Ablation Experiment of Change Detection Unsupervised Domain Adaptation Model (CD-UDA)

According to the DA strategy and the sample selection strategy introduced in Section 3.2, the effectiveness of each DA strategy is verified by an ablation experiment. This experiment has the following three parts:

1. D0: Evaluate the UDA performance of the model when pre-trained with the source domain data, using only the mid-layer marginal distribution alignment (Section 3.2.1);
2. D1: Evaluate the UDA performance of the model when pre-trained with the source domain data, using transferable features derived from the probabilistic easy-to-hard target domain sample selection strategy (Section 3.2.3) for the inter-domain marginal distribution alignment (Section 3.2.1) prior to the classification layer;
3. D2: Building on D1, evaluate the UDA performance of the mid-layer marginal distribution alignment (Section 3.2.2).

The results of the ablation experiments are shown in Table 3, where $G \rightarrow L$ refers to the source domain data being GZ and the target domain data being LEVIR. These

results show that the proposed mid-level domain edge distribution domain alignment, deep conditional distribution domain alignment, and sample selection strategies mainly improve the performance of CD-UDA. Among them, the probability-based easy-to-hard sample selection strategy significantly mitigates the convergence of CD-UDA to a local optimum, due to sample imbalance. The D1 approach with conditional distribution domain constraints shows significant overall improvement in the experiments, and in particular, the UDA strategy is more effective in mitigating the severe sample imbalance in the CD task. The above experimental results show that the sample imbalance problem can be alleviated by setting the same number of transferable samples for the changed and unchanged regions in the sample screening process. However, if the mid-level marginal distribution adaptation is not included, the convergence direction of the model will be different from that of the source domain model, i.e., the prior knowledge of the source domain is ignored. Given that the target domain data are unlabeled, CD-UDA can only improve the target domain test performance by minimizing the feature distribution difference between the source and target domain data.

Table 3. Ablation experimental results for domain adaptation strategies with different source domains.

D0	D1	D2	Data	OA	UAcc	CAcc	F1	mIoU
✓	×	×	$G \rightarrow L$	0.973	0.992	0.232	0.643	0.568
✓	✓	×	$G \rightarrow L$	0.964	0.989	0.257	0.646	0.577
✓	✓	✓	$G \rightarrow L$	0.974	0.990	0.305	0.673	0.597
✓	×	×	$G \rightarrow S$	0.796	0.992	0.076	0.511	0.434
✓	✓	×	$G \rightarrow S$	0.792	0.984	0.087	0.517	0.435
✓	✓	✓	$G \rightarrow S$	0.797	0.966	0.188	0.584	0.478
✓	×	×	$L \rightarrow S$	0.786	0.976	0.088	0.514	0.432
✓	✓	×	$L \rightarrow S$	0.785	0.981	0.065	0.496	0.421
✓	✓	✓	$L \rightarrow S$	0.777	0.958	0.111	0.524	0.434
✓	×	×	$L \rightarrow G$	0.941	0.994	0.138	0.597	0.534
✓	✓	×	$L \rightarrow G$	0.941	0.987	0.247	0.657	0.574
✓	✓	✓	$L \rightarrow G$	0.940	0.978	0.361	0.698	0.606
✓	×	×	$S \rightarrow L$	0.939	0.953	0.333	0.587	0.521
✓	✓	×	$S \rightarrow L$	0.843	0.845	0.769	0.551	0.472
✓	✓	✓	$S \rightarrow L$	0.951	0.966	0.348	0.614	0.548
✓	×	×	$S \rightarrow G$	0.671	0.654	0.924	0.524	0.400
✓	✓	×	$S \rightarrow G$	0.852	0.856	0.788	0.657	0.547
✓	✓	✓	$S \rightarrow G$	0.877	0.886	0.733	0.678	0.571

There are two customized parameters, T_p and v , in the sample screening strategy Algorithm 2, and the specific values of the parameters are determined through ablation experiments. This experiment is carried out on the basis of all the additions to the D2 CD-UDA strategy. The ablation experiment is first performed for the initial probability threshold $\{T_p | 0.5 < T_p \leq 1\}$. Not performing the second screening will lead to model optimization failure due to sample imbalance, so here the second screening parameter v is set to the empirical value of 0.9. Table 4 shows the performance of $T_p \in \{0.6, 0.7, 0.8, 0.9\}$ on different cross-sample datasets, where the experimental metrics are $F1$ and $mIOU$. After determining the value of $T_p = 0.9$, ablation experiments are performed on the second sample screening weight v . Table 5 shows the performance of $v \in \{0.6, 0.7, 0.8, 0.9\}$ on different cross-sample datasets where the experimental metrics are $F1$ and $mIOU$. Among them, the bolded result is the best performance. From the results of the experiment, it was determined that $v = 0.8$.

Table 4. Effect of the initial probability threshold T_p ($F1/mIOU$).

T_p	0.6	0.7	0.8	0.9
$G \rightarrow L$	0.534/0.509	0.608/0.549	0.613/0.608	0.615/0.556
$G \rightarrow S$	0.497/0.420	0.558/0.522	0.562/0.531	0.568/0.511
$L \rightarrow S$	0.488/0.418	0.485/0.417	0.502/0.419	0.497/0.418
$L \rightarrow G$	0.613/0.544	0.643/0.565	0.654/0.573	0.660/0.581
$S \rightarrow L$	0.542/0.499	0.579/0.531	0.582/0.532	0.610/0.553
$S \rightarrow G$	0.533/0.495	0.594/0.543	0.597/0.548	0.604/0.493

Table 5. Effect of the initial probability threshold v ($F1/mIOU$).

v	0.6	0.7	0.8	0.9
$G \rightarrow L$	0.583/0.536	0.655/0.580	0.673/0.597	0.658/0.584
$G \rightarrow S$	0.535/0.456	0.557/0.478	0.584/0.479	0.542/0.461
$L \rightarrow S$	0.485/0.416	0.497/0.418	0.523/0.434	0.510/0.428
$L \rightarrow G$	0.627/0.564	0.684/0.652	0.698/0.610	0.661/0.585
$S \rightarrow L$	0.596/0.529	0.615/0.551	0.614/0.548	0.614/0.549
$S \rightarrow G$	0.608/0.549	0.627/0.564	0.659/0.562	0.627/0.564

4.5. Performance Evaluation of Change Detection Unsupervised Domain Adaptation Model (CD-UDA)

The research on CD-UDA is very limited. In order to evaluate the performance of the proposed CD-UDA, the experiments are divided into the following three parts: first, evaluate the transferable characteristics of the multi-domain fully-connected layer by MK-MMD using the DSDANet model; second, incorporate the inter-domain conditional distribution difference and probabilistic sample-selection-based transfer strategy proposed in this paper into the two-stream UCDNet and ISNet models, so as to verify the generality of the CD-UDA method proposed in this paper; third, generalize the proposed CD-UDA method by creating a benchmark; select several CD models without a DA strategy but with better performance for cross-domain dataset testing. The experimental results are shown in Tables 6–11, with the best results in bold.

UDA is essentially a migration of existing knowledge from the source domain, and it is impossible to have a situation where knowledge is created out of nothing. Therefore, it is first necessary to analyze the situation of the dataset to be migrated, and then obtain the application scenario of CD-UDA. From the experimental results and the analysis of the experimental process, it is known that the CD results, without adding the UDA strategy, can be used to obtain the distribution relationship between the datasets, using the test results on the target domain, although the experimental results, in this case, may have drastic oscillations within the iteration process. Based on the experimental results of the model without the CD-UDA strategy, the following conclusions can be drawn: firstly, the C_{Acc} metrics are significantly lower in most experimental results, which indicates richer content information and greater domain uncertainty in the change region in the CD task; secondly, $F1$ and C_{Acc} are higher in $S \rightarrow L$ (Table 10) and $S \rightarrow G$ (Table 11), which can be attributed to the fact that SUYU and the other two datasets differ not only in the distribution of the data, but also in the range of the labeling—the labeling range of SUYU is larger than the other two datasets; thirdly, as shown in Tables 6 and 9, the domain similarity is more obvious in the G and L datasets.

Table 6. Performance of CD-UDA with the GZ source domain and the LEVIR target domain (GZ→LEVIR).

Method	OA	UAcc	CAcc	F1	mIoU
DeepLab	0.970	0.992	0.095	0.559	0.521
FCSiamD	0.973	0.994	0.087	0.551	0.522
EUNet	0.976	0.999	0.030	0.522	0.502
UCDNet	0.974	0.997	0.044	0.460	0.507
ISNet	0.969	0.988	0.218	0.621	0.558
IRDNet	0.974	0.991	0.176	0.609	0.552
IRD-CD-UDA-Source	0.969	0.988	0.930	0.899	0.909
UCDNet-UDA	0.966	0.998	0.052	0.541	0.513
FCSiamD-UDA	0.966	0.988	0.074	0.541	0.509
DSDANet	0.941	0.994	0.126	0.589	0.528
IRD-CD-UDA	0.974	0.990	0.305	0.673	0.597

Table 7. Performance of CD-UDA with the GZ source domain and the SUYU target domain (GZ→SUYU).

Method	OA	UAcc	CAcc	F1	mIoU
DeepLab	0.794	0.996	0.053	0.491	0.422
FCSiamD	0.793	0.995	0.054	0.508	0.422
EUNet	0.790	0.993	0.044	0.482	0.415
UCDNet	0.789	0.998	0.020	0.532	0.404
ISNet	0.786	0.979	0.077	0.505	0.427
IRDNet	0.789	0.980	0.081	0.510	0.430
IRD-CD-UDA-Source	0.979	0.990	0.927	0.901	0.901
UCDNet-UDA	0.791	0.977	0.101	0.527	0.440
FCSiamD-UDA	0.785	0.991	0.039	0.485	0.408
DSDANet	0.788	0.992	0.039	0.477	0.412
IRD-CD-UDA	0.800	0.966	0.188	0.585	0.478

From the experimental results, it can be seen that the DSDANet model for the CD-DA task has more obvious instability, which is caused by the following two reasons: firstly, due to the sample imbalance in the CD dataset, if a suitable sample selection strategy is not adopted, it will lead to the model easily converging to the significant change region and the unchanged region, which will produce negative migration; secondly, due to the specificity of the CD dual-stream model, if domain feature alignment is only implemented in the classification layer, the style information of the samples will be ignored, and the aligned domain features only characterize the change detection results.

In order to evaluate the generalization of the CD-UDA strategy proposed in this paper, FCSiamD and UCDNet are chosen as the experiments for the addition of the CD-UDA strategy. Since FCSiamD and UCDNet cannot generate transferable features applicable to the marginal distribution alignment in the middle layer, only the conditional distribution alignment strategy and the sample selection strategy for the classification layer are added to the experiment. Overall, the addition of the proposed UDA strategy to the CD model results in more significant improvements, especially in S→L (Table 10) and S→G (Table 11).

From L→S (Table 8) and G→S (Table 7), the results have high accuracy for unchanged regions, low accuracy for changed regions, and insignificant CD-UDA performance. Combined with the overview of the experimental data (Table 1), we believe that this problem is caused by the following two problems: firstly, the difference between SUYU and the other two datasets not only in data distribution, but also in the labeling range (discrimination threshold for the changed region), which is larger in SUYU than in the other two

datasets; and, secondly, the source domain data with a serious sample imbalance increase the difficulty of performing the CD-UDA task.

Table 8. Performance of CD-UDA with the LEVIR source domain and the SUYU target domain (LEVIR → SUYU).

Method	OA	UAcc	CAcc	F1	mIoU
DeepLab	0.787	0.999	0.005	0.445	0.395
FCSiamD	0.788	0.999	0.011	0.451	0.399
EUNet	0.787	0.999	0.005	0.445	0.396
UCDNet	0.787	0.999	0.006	0.447	0.396
ISNet	0.789	0.997	0.023	0.462	0.405
IRDNet	0.786	0.999	0.004	0.444	0.395
IRD-CD-UDA-Source	0.989	0.991	0.851	0.928	0.881
UCDNet-UDA	0.792	0.997	0.040	0.479	0.415
FCSiamD-UDA	0.786	0.999	0.046	0.453	0.405
DSDANet	0.786	0.999	0.054	0.441	0.399
IRD-CD-UDA	0.787	0.958	0.112	0.524	0.434

Table 9. Performance of CD-UDA with the LEVIR source domain and the GZ target domain (LEVIR → GZ).

Method	OA	UAcc	CAcc	F1	mIoU
DeepLab	0.933	0.978	0.258	0.613	0.564
FCSiamD	0.942	0.994	0.152	0.608	0.541
EUNet	0.928	0.967	0.338	0.665	0.576
UCDNet	0.941	0.994	0.131	0.593	0.531
ISNet	0.913	0.988	0.228	0.648	0.621
IRD-CD-UDA-Source	0.991	0.996	0.851	0.938	0.888
IRDNet	0.948	0.992	0.299	0.613	0.607
IRD-CD-UDA-Source	0.989	0.991	0.851	0.928	0.881
UCDNet-UDA	0.912	0.953	0.251	0.605	0.528
FCSiamD-UDA	0.917	0.959	0.288	0.629	0.547
DSDANet	0.927	0.968	0.304	0.652	0.566
IRD-CD-UDA	0.950	0.980	0.361	0.698	0.610

Table 10. Performance of CD-UDA with the SUYU source domain and the LEVIR target domain (SUYU → LEVIR).

Method	OA	UAcc	CAcc	F1	mIoU
DeepLab	0.947	0.966	0.177	0.556	0.511
FCSiamD	0.950	0.969	0.207	0.571	0.521
EUNet	0.754	0.755	0.733	0.491	0.408
UCDNet	0.600	0.601	0.552	0.373	0.313
ISNet	0.847	0.851	0.676	0.574	0.470
IRDNet	0.875	0.880	0.664	0.567	0.493
IRD-CD-UDA-Source	0.901	0.958	0.691	0.844	0.749
UCDNet-UDA	0.901	0.922	0.272	0.537	0.486
FCSiamD-UDA	0.969	0.992	0.103	0.533	0.512
DSDANet	0.874	0.877	0.769	0.578	0.499
IRD-CD-UDA	0.951	0.966	0.348	0.614	0.548

Table 11. Performance of CD-UDA with the SUYU source domain and the GZ target domain (SUYU→GZ).

Method	OA	UAcc	CAcc	F1	mIoU
DeepLab	0.800	0.800	0.792	0.596	0.493
FCSiamD	0.880	0.906	0.496	0.637	0.541
EUNet	0.536	0.508	0.957	0.438	0.310
UCDNet	0.752	0.766	0.537	0.461	0.431
ISNet	0.807	0.812	0.730	0.604	0.494
IRDNet	0.815	0.816	0.789	0.619	0.507
IRD-CD-UDA-Source	0.902	0.951	0.681	0.841	0.742
UCDNet-UDA	0.818	0.821	0.766	0.619	0.508
FCSiamD-UDA	0.877	0.886	0.733	0.668	0.570
DSDANet	0.781	0.777	0.843	0.597	0.481
IRD-CD-UDA	0.898	0.924	0.491	0.659	0.562

From the experimental results of S→L (Table 10) and S→G (Table 11), we can see that most of the methods improve their results significantly after adding the proposed UDA strategy. Combined with the overview of the experimental data (Table 1), we believe that the performance of the CD-UDA model on the target domain dataset is enhanced when the source domain data are broader and the feature distribution is wider.

From the experimental results of G→L (Table 6) and L→G (Table 9), it can be seen that CD-UDA mitigates the sample imbalance, resulting in the easier convergence of the target domain test results with the unchanged region. Combined with the overview of the experimental data (Table 1), CD-UDA is more effective for datasets with similar labeling scales (GZ and LEVIR) and is also more stable during the training process.

The proposed IRD-CD-UDA achieves the best performance in most of the cross-domain data experiments, and in particular, the performance of the F1 and mIoU metrics is significantly improved, with F1 improving by 3–22% and Miou by 2–13%. This demonstrates the effectiveness of the proposed CD-UDA in mitigating the sample imbalance problem, which often causes the model to converge to a local minimum. Meanwhile, the OA performance achieves the best results on G→S, L→S, L→G, and S→G, and outperforms most of the comparison methods on S→L and L→S. The results show that the CD-UDA proposed in this paper improves the performance of the target domain data without destroying the prior source domain and improves the generalization ability of the model.

4.6. Visualization

A set of bi-temporal data randomly selected from the three datasets are fed into the trained IRD-CD-UDA to produce visualization results. As shown in Figure 4, the experimental results indicate that the annotation criteria for the CD dataset vary across the datasets. These criteria dictate the designation of regions as changed or unchanged, which directly affects the performance of the CD-UDA model. As shown in Figure 4, the model exhibits increased sensitivity to regional changes when the source domain is SUYU, compared to the LEVIR and GZ datasets. In Figure 4, due to the large amount of data in the SUYU dataset and the richness of annotation categories, the changed regions of the GZ data can be better identified.

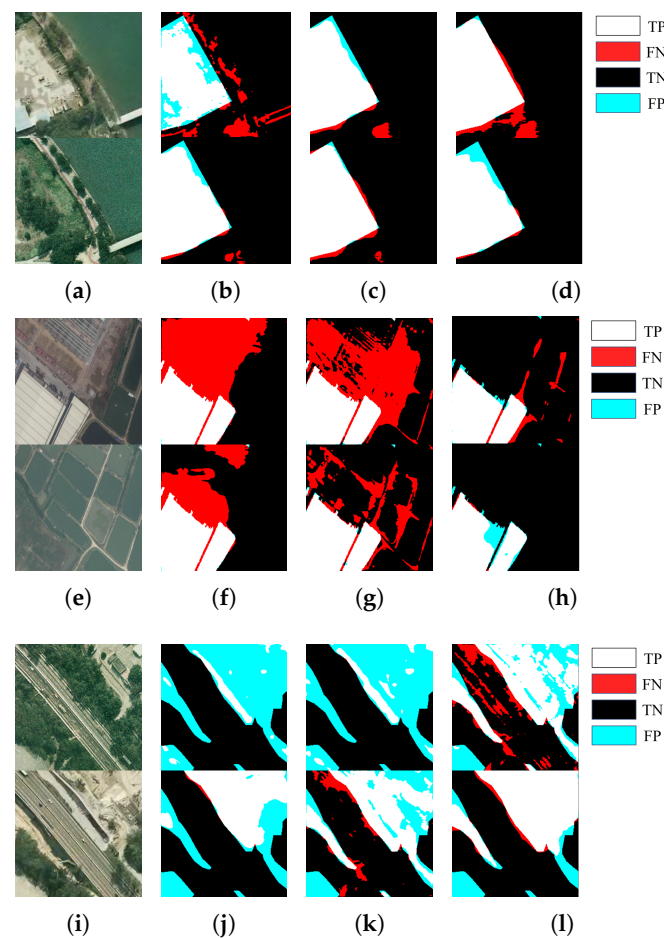


Figure 4. CD-UDA visualization results, where black is TN, white is TP, blue is FP, and red is FN. TN indicates that the unchanged pixel is correctly detected, TP indicates that the changed pixel is correctly detected, FN indicates the unchanged pixel is incorrectly detected, and FP indicates the changed pixel is incorrectly detected. (a–d) are LEVIR→SYSU (top) and GZ→SYSU (bottom) experimental results, and the UDA methods are, respectively, FCSiamD-UDA, DSDANet, and IRD-CD-UDA; (e–h) are LEVIR→GZ (top) and SYSU→GZ (bottom) experimental results, and the UDA methods are, respectively, FCSiamD-UDA, DSDANet, and MDCS-CD-UDA; (i–l) are SYSU→LEVIR (top) and GZ→LEVIR (bottom) experimental results, and the UDA methods are, respectively, FCSiamD-UDA, DSDANet, and IRD-CD-UDA.

To visualize the distribution of differential features before and after CD-UDA, the T-SNE algorithm [51] was used to plot the classification feature distribution before and after domain adaptation, as shown in Figure 5, which shows the altered/invariant region-coupled features, altered intra-class features, and modified intra-class features. Obviously, prior to DA (Figure 4a–c), the distributions of the coupled features and the intra-class features are mixed, with the intra-class feature distributions being distinctly different for different domains. After DA (Figure 5d–f), the distributions of the coupled features are clustered by distinct classes, while the intra-class feature distributions of different domains approximate by class. Due to the sample imbalance, the model exhibits reduced confidence in the features of changing regions within a few target domains, leading to the conflation of depth features of these changing regions with features of other categories.

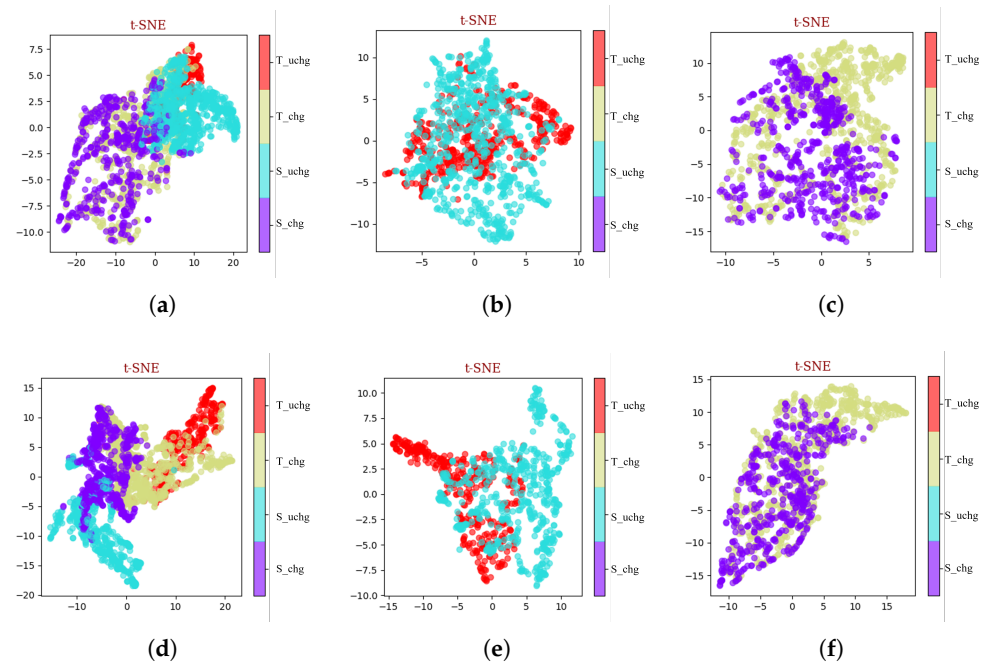


Figure 5. Feature distribution before and after DA. $x_{t,unchg}$ is the transferable unchanged feature vector of a point in the target domain, $x_{t,unchg} \in x_t$; $x_{t,chg}$ is the transferable changed feature vector of a point in the target domain, $x_{t,chg} \in x_t$; $x_{s,unchg}$ is the transferable unchanged feature vector of a point in the source domain, $x_{s,unchg} \in x_s$; and $x_{s,chg}$ is the transferable changed feature vector of a point in the source domain, $x_{s,chg} \in x_s$. (a–c) Feature distribution visualization before DA. (d–f) Feature distribution visualization after DA.

5. Conclusions

In this paper, a multi-scale unsupervised domain adaptation model for illumination–reflection decoupled change detection (IRD-CD-UDA) is proposed. The illumination–reflection decoupled CD model is able to extract and fuse illumination–reflection features to improve the supervised CD performance. Three strategies are designed to address the specific characteristics of CD-UDA: a domain alignment method for the marginal distribution of global features in the mid-layer, a domain alignment scheme for the conditional distribution of features in the classification layer, and an easy-to-hard sample selection method based on the probabilistic threshold of the CD map. The experiments verify that the proposed IRD-CD-UDA model can effectively improve the performance of a new homogeneous CD dataset (target domain) without additional data labeling. In addition, by performing ablation experiments and comparative analysis on three different CD datasets, it is confirmed that embedding the proposed CD-UDA strategy into different CD models can effectively improve their performance on different datasets, and ultimately improve the reusability of CD models.

Author Contributions: Conceptualization, J.Y. and R.F.; methodology, R.F.; software, R.F.; validation, R.F. and J.X.; formal analysis, R.F.; resources, J.Y. and Z.H.; data curation, Y.X.; writing—original draft preparation, R.F.; writing—review and editing, J.Y. and H.H.; visualization, R.F.; supervision, J.X.; project administration, J.Y.; funding acquisition, J.Y. and H.H. All authors have read and agreed to the published version of the manuscript.

Funding: This work was financially supported by the National Natural Science Foundation of China (Grant Nos. 12174314) and Innovation Foundation for Doctor Dissertation of Northwestern Polytechnical University Nos. CX2023064.

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

Abbreviations

The following abbreviations are used in this manuscript:

CD	Change Detection
UDA	Unsupervised Domain Adaptation
IRD	Illumination-Reflection Feature Decoupling
ASSP	Atrous Apatial Pyramid Pooling

References

- Shi, W.; Zhang, M.; Zhang, R.; Chen, S.; Zhan, Z. Change Detection Based on Artificial Intelligence: State-of-the-Art and Challenges. *Remote Sens.* **2020**, *12*, 1688. [\[CrossRef\]](#)
- Zhang, R.; Zhang, H.; Ning, X.; Huang, X.; Wang, J.; Cui, W. Global-aware siamese network for change detection on remote sensing images. *ISPRS J. Photogramm. Remote Sens.* **2023**, *199*, 61–72. [\[CrossRef\]](#)
- Wang, X.; Yan, X.; Tan, K.; Pan, C.; Ding, J.; Liu, Z.; Dong, X. Double U-Net (W-Net): A change detection network with two heads for remote sensing imagery. *Int. J. Appl. Earth Obs. Geoinf.* **2023**, *122*, 103456. [\[CrossRef\]](#)
- Li, S.; Wang, Y.; Cai, H.; Lin, Y.; Wang, M.; Teng, F. MF-SRCDNet: Multi-feature fusion super-resolution building change detection framework for multi-sensor high-resolution remote sensing imagery. *Int. J. Appl. Earth Obs. Geoinf.* **2023**, *119*, 103303. [\[CrossRef\]](#)
- Kumar, A.; Mishra, V.; Panigrahi, R.K.; Martorella, M. Application of Hybrid-Pol SAR in Oil-Spill Detection. *IEEE Geosci. Remote Sens. Lett.* **2023**, *20*, 1–5. [\[CrossRef\]](#)
- Ioffe, S.; Szegedy, C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In Proceedings of the International Conference on Machine Learning, Lille, France, 6–11 July 2015; pp. 448–456. [\[CrossRef\]](#)
- Wu, C.; Du, B.; Zhang, L. Slow Feature Analysis for Change Detection in Multispectral Imagery. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 2858–2874. [\[CrossRef\]](#)
- Li, H.C.; Yang, G.; Yang, W.; Du, Q.; Emery, W.J. Deep nonsmooth nonnegative matrix factorization network with semi-supervised learning for SAR image change detection. *ISPRS J. Photogramm. Remote Sens.* **2020**, *160*, 167–179. [\[CrossRef\]](#)
- Zhang, X.; Su, H.; Zhang, C.; Gu, X.; Tan, X.; Atkinson, P.M. Robust unsupervised small area change detection from SAR imagery using deep learning. *ISPRS J. Photogramm. Remote Sens.* **2021**, *173*, 79–94. [\[CrossRef\]](#)
- Jiang, X.; Li, G.; Zhang, X.P.; He, Y. A Semisupervised Siamese Network for Efficient Change Detection in Heterogeneous Remote Sensing Images. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 1–18. [\[CrossRef\]](#)
- Xie, S.; Zheng, Z.; Chen, L.; Chen, C. Learning Semantic Representations for Unsupervised Domain Adaptation. In Proceedings of the 35th International Conference on Machine Learning PMLR, Stockholm, Sweden, 10–15 July 2018; Dy, J., Krause, A., Eds.; Volume 80, pp. 5423–5432.
- 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. [\[CrossRef\]](#)
- Li, X.; Du, Z.; Huang, Y.; Tan, Z. A deep translation (GAN) based change detection network for optical and SAR remote sensing images. *ISPRS J. Photogramm. Remote Sens.* **2021**, *179*, 14–34. [\[CrossRef\]](#)
- Zhao, W.; Chen, X.; Ge, X.; Chen, J. Using Adversarial Network for Multiple Change Detection in Bitemporal Remote Sensing Imagery. *IEEE Geosci. Remote Sens. Lett.* **2022**, *19*, 1–5. [\[CrossRef\]](#)
- Li, J.; Zi, S.; Song, R.; Li, Y.; Hu, Y.; Du, Q. A Stepwise Domain Adaptive Segmentation Network With Covariate Shift Alleviation for Remote Sensing Imagery. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 1–15. [\[CrossRef\]](#)
- Long, M.; Cao, Y.; Wang, J.; Jordan, M. Learning transferable features with deep adaptation networks. In Proceedings of the International Conference on Machine Learning, PMLR, Lille, France, 7–9 July 2015; pp. 97–105. [\[CrossRef\]](#)
- Geng, J.; Deng, X.; Ma, X.; Jiang, W. Transfer Learning for SAR Image Classification Via Deep Joint Distribution Adaptation Networks. *IEEE Trans. Geosci. Remote Sens.* **2020**, *58*, 5377–5392. [\[CrossRef\]](#)
- Zhao, H.; Combes, R.; Zhang, K.; Gordon, G.J. On Learning Invariant Representation for Domain Adaptation. *arXiv* **2019**, arXiv:1901.09453.
- Pan, S.J.; Tsang, I.W.; Kwok, J.T.; Yang, Q. Domain Adaptation via Transfer Component Analysis. *IEEE Trans. Neural Netw.* **2011**, *22*, 199–210. [\[CrossRef\]](#)
- Zellinger, W.; Grubinger, T.; Lughofer, E.; Natschläger, T.; Saminger-Platz, S. Central moment discrepancy (cmd) for domain-invariant representation learning. *arXiv* **2017**, arXiv:1702.08811.
- Sun, B.; Saenko, K. Deep coral: Correlation alignment for deep domain adaptation. In Proceedings of the Computer Vision–ECCV 2016 Workshops: Amsterdam, The Netherlands, 8–10 and 15–16 October 2016; Proceedings, Part III 14; Springer: Berlin/Heidelberg, Germany, 2016; pp. 443–450. [\[CrossRef\]](#)
- Gretton, A.; Sejdinovic, D.; Strathmann, H.; Balakrishnan, S.; Pontil, M.; Fukumizu, K.; Sriperumbudur, B.K. Optimal kernel choice for large-scale two-sample tests. *Adv. Neural Inf. Process. Syst.* **2012**, *25*. [\[CrossRef\]](#)
- Long, M.; Wang, J.; Ding, G.; Sun, J.; Yu, P.S. Transfer Feature Learning with Joint Distribution Adaptation. In Proceedings of the 2013 IEEE International Conference on Computer Vision (CVPR), Portland, OR, USA, 23–28 June 2013; pp. 2200–2207. [\[CrossRef\]](#)

24. Zhang, J.; Li, W.; Ogunbona, P. Joint Geometrical and Statistical Alignment for Visual Domain Adaptation. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 5150–5158. [\[CrossRef\]](#)
25. Tzeng, E.; Hoffman, J.; Zhang, N.; Saenko, K.; Darrell, T. Deep domain confusion: Maximizing for domain invariance. *arXiv* **2014**, arXiv:1412.3474.
26. Long, M.; Zhu, H.; Wang, J.; Jordan, M.I. Deep transfer learning with joint adaptation networks. In Proceedings of the International Conference on Machine Learning, Sydney, Australia, 6–11 August 2017; pp. 2208–2217. [\[CrossRef\]](#)
27. Wang, J.; Chen, Y.; Hao, S.; Feng, W.; Shen, Z. Balanced Distribution Adaptation for Transfer Learning. In Proceedings of the 2017 IEEE International Conference on Data Mining (ICDM), New Orleans, LA, USA, 18–21 November 2017; pp. 1129–1134. [\[CrossRef\]](#)
28. Wang, J.; Chen, Y.; Feng, W.; Yu, H.; Huang, M.; Yang, Q. Transfer learning with dynamic distribution adaptation. *ACM Trans. Intell. Syst. Technol. (TIST)* **2020**, *11*, 1–25. [\[CrossRef\]](#)
29. Yan, H.; Ding, Y.; Li, P.; Wang, Q.; Xu, Y.; Zuo, W. Mind the Class Weight Bias: Weighted Maximum Mean Discrepancy for Unsupervised Domain Adaptation. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 945–954. [\[CrossRef\]](#)
30. Shen, J.; Qu, Y.; Zhang, W.; Yu, Y. Wasserstein distance guided representation learning for domain adaptation. In Proceedings of the AAAI Conference on Artificial Intelligence, New Orleans, LA, USA, 2–7 February 2018; Volume 32. [\[CrossRef\]](#)
31. Kang, G.; Jiang, L.; Yang, Y.; Hauptmann, A.G. Contrastive Adaptation Network for Unsupervised Domain Adaptation. In Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 15–20 June 2019; pp. 4888–4897. [\[CrossRef\]](#)
32. Chen, H.; Wu, C.; Du, B.; Zhang, L. DSDANet: Deep Siamese domain adaptation convolutional neural network for cross-domain change detection. *arXiv* **2020**, arXiv:2006.09225v1.
33. Ahmed, S.M.; Raychaudhuri, D.S.; Paul, S.; Oymak, S.; Roy-Chowdhury, A.K. Unsupervised Multi-source Domain Adaptation Without Access to Source Data. In Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, TN, USA, 20–25 June 2021; pp. 10098–10107. [\[CrossRef\]](#)
34. Liang, J.; Hu, D.; Feng, J. Domain Adaptation with Auxiliary Target Domain-Oriented Classifier. In Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, TN, USA, 20–25 June 2021; pp. 16627–16637. [\[CrossRef\]](#)
35. Yang, L.; Balaji, Y.; Lim, S.N.; Shrivastava, A. Curriculum manager for source selection in multi-source domain adaptation. In Proceedings of the Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, 23–28 August 2020; Proceedings, Part XIV 16; Springer: Berlin/Heidelberg, Germany, 2020; pp. 608–624. [\[CrossRef\]](#)
36. Chen, L.C.; Papandreou, G.; Kokkinos, I.; Murphy, K.; Yuille, A.L. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. *IEEE Trans. Pattern Anal. Mach. Intell.* **2018**, *40*, 834–848. [\[CrossRef\]](#) [\[PubMed\]](#)
37. Wu, H.; Zheng, S.; Zhang, J.; Huang, K. Fast end-to-end trainable guided filter. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 1838–1847. [\[CrossRef\]](#)
38. Weyermann, J.; Kneubühler, M.; Schläpfer, D.; Schaepman, M.E. Minimizing Reflectance Anisotropy Effects in Airborne Spectroscopy Data Using Ross–Li Model Inversion With Continuous Field Land Cover Stratification. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 5814–5823. [\[CrossRef\]](#)
39. Petro, A.B.; Sbert, C.; Morel, J.M. Multiscale Retinex. *Image Process. Line* **2014**, 71–88. [\[CrossRef\]](#)
40. Hu, J.; Shen, L.; Sun, G. Squeeze-and-Excitation Networks. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 7132–7141. [\[CrossRef\]](#)
41. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the 2016 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778. [\[CrossRef\]](#)
42. Shi, Q.; Liu, M.; Li, S.; Liu, X.; Wang, F.; Zhang, L. A Deeply Supervised Attention Metric-Based Network and an Open Aerial Image Dataset for Remote Sensing Change Detection. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 1–16. [\[CrossRef\]](#)
43. Peng, D.; Bruzzone, L.; Zhang, Y.; Guan, H.; Ding, H.; Huang, X. SemiCDNet: A Semisupervised Convolutional Neural Network for Change Detection in High Resolution Remote-Sensing Images. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 5891–5906. [\[CrossRef\]](#)
44. Chen, H.; Shi, Z. A spatial-temporal attention-based method and a new dataset for remote sensing image change detection. *Remote Sens.* **2020**, *12*, 1662. [\[CrossRef\]](#)
45. Xie, B.; Yuan, L.; Li, S.; Liu, C.H.; Cheng, X. Towards Fewer Annotations: Active Learning via Region Impurity and Prediction Uncertainty for Domain Adaptive Semantic Segmentation. In Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), New Orleans, LA, USA, 18–24 June 2022; pp. 8058–8068. [\[CrossRef\]](#)
46. Luo, X.; Li, X.; Wu, Y.; Hou, W.; Wang, M.; Jin, Y.; Xu, W. Research on Change Detection Method of High-Resolution Remote Sensing Images Based on Subpixel Convolution. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 1447–1457. [\[CrossRef\]](#)
47. Caye Daudt, R.; Le Saux, B.; Boulch, A. Fully Convolutional Siamese Networks for Change Detection. In Proceedings of the 2018 25th IEEE International Conference on Image Processing (ICIP), Athens, Greece, 7–10 October 2018; pp. 4063–4067. [\[CrossRef\]](#)
48. Raza, A.; Huo, H.; Fang, T. EUNet-CD: Efficient UNet++ for Change Detection of Very High-Resolution Remote Sensing Images. *IEEE Geosci. Remote Sens. Lett.* **2022**, *19*, 1–5. [\[CrossRef\]](#)

-
49. Basavaraju, K.S.; Sravya, N.; Lal, S.; Nalini, J.; Reddy, C.S.; Acqua, D. UCDNet: A Deep Learning Model for Urban Change Detection From Bi-Temporal Multispectral Sentinel-2 Satellite Images. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 1–10. [[CrossRef](#)]
 50. Cheng, G.; Wang, G.; Han, J. ISNet: Towards Improving Separability for Remote Sensing Image Change Detection. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 1–11. [[CrossRef](#)]
 51. Maaten, L. Accelerating t-SNE using tree-based algorithms. *J. Mach. Learn. Res.* **2014**, *15*, 3221–3245. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.