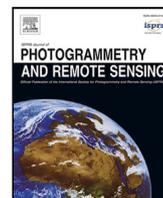




Contents lists available at ScienceDirect

## ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: [www.elsevier.com/locate/isprsjprs](http://www.elsevier.com/locate/isprsjprs)

# AWDA: Adversarial and Weighted Domain Adaptation for cross-dataset change detection

Xueting Zhang, Xin Huang<sup>✉</sup>\*, Jiayi Li

School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, 430079, China

## ARTICLE INFO

### Keywords:

Change detection  
Adversarial learning  
Class weighting  
Domain adaptation  
Transfer learning

## ABSTRACT

Recent advancements in change detection (CD) using fully-supervised methods have been significant; however, effectively applying CD in scenarios where labels are unavailable remains a challenge. To address this, our study introduces a new task, domain adaptive change detection (DACD), which transfers change knowledge from a labeled CD dataset (source domain) to an unlabeled CD dataset (target domain). In practice, two challenges hinder change knowledge transfer across domains: domain shifts, such as resolution differences and change semantic discrepancies, and imbalanced distribution between the minority change class and the dominant no-change class. To tackle these issues, we propose a novel Adversarial and Weighted Domain Adaptation (AWDA) framework for DACD. AWDA employs a Siamese encoder–decoder network shared between source and target domains to extract features and make predictions from bi-temporal remote sensing images. Moreover, AWDA incorporates three cross-domain learning strategies for learning domain-invariant CD representations: (1) supervised learning, which uses all the labeled data of the source domain to train the model to obtain initial CD capability, (2) domain adversarial training, which aligns the features between the source and target domains adversarially, and (3) class-weighted self-training, which dynamically computes and assigns class weights for the self-training on the unlabeled data of the target domain. The proposed AWDA effectively mitigates cross-domain shifts and preserves the integrity of the minor change class during knowledge transfer. To evaluate our method's effectiveness, we conducted comprehensive experiments across four cross-domain CD scenarios using three well-known building CD datasets. The results demonstrate AWDA substantially enhances CD performance in the target domain, achieving IoU increase ranging from 13.64 to 34.73, and significantly surpassing several competing domain adaptation methods. Our code will be available at <https://github.com/zxt9/AWDA>.

## 1. Introduction

As a fundamental task in Earth observation, change detection (CD) (Zhang et al., 2023b,a) is pivotal in analyzing changes on the Earth's surface over different periods, using bi-temporal remote sensing image (RSI) pairs. It has a broad downstream applications in environmental monitoring (Alesheikh et al., 2007; Liu et al., 2019; Decuyper et al., 2022), resource management (Kennedy et al., 2009; Ngondo et al., 2021), and disaster assessment (Giustarini et al., 2012; Anniballe et al., 2018; Qing et al., 2022). Deep learning-based methods, reliant on extensive annotated training data, have marked a significant advance in CD tasks (Zhan et al., 2017; Mou et al., 2018; Wang et al., 2018; Shafique et al., 2022; Bai et al., 2023; Dong et al., 2024) due to their promising performance. However, the annotation of training data often requires substantial labor and time, posing challenges in practical scenarios with limited or no labels.

To address this issue, our study introduces the concept of domain adaptation (DA) to the CD scenarios, termed as Domain Adaptive Change Detection (DACD) as shown in Fig. 1. DACD focuses on transferring change knowledge from a labeled CD dataset (source domain) where annotations are fully provided, to an unlabeled CD dataset (target domain) where labels are inaccessible. This task needs to bridge the gap between the two domains, leveraging the information from the labeled source domain to enhance CD performance in the unlabeled target domain.

However, two key challenges hinder the cross-domain change knowledge transfer: (1) domain shifts between the two domains, such as various imaging sensors, different spatial resolutions, and diverse change semantics, and (2) imbalanced distribution between the non-change and change classes, with non-change areas predominantly occupying the major pixels of RSI pairs and change areas comprising

\* Corresponding author.

E-mail address: [xhuang@whu.edu.cn](mailto:xhuang@whu.edu.cn) (X. Huang).

<https://doi.org/10.1016/j.isprsjprs.2025.04.008>

Received 17 July 2024; Received in revised form 12 January 2025; Accepted 2 April 2025

Available online 22 April 2025

0924-2716/© 2025 International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS). Published by Elsevier B.V. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

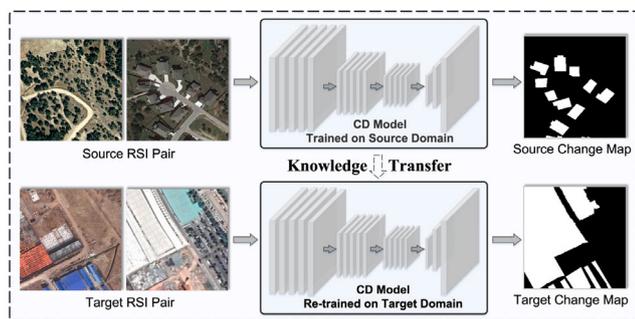


Fig. 1. The illustration of DACD. The CD model is trained on a labeled source domain and transferred to an unlabeled target domain.

only the remaining small pixels. Both the issues significantly degrade the CD model's ability to directly detect the change areas in the target domain.

To address the challenges, this study proposes a novel DACD framework termed Adversarial and Weighted Domain Adaptation (AWDA). AWDA employs a Siamese encoder–decoder network shared between the source and target domains to extract features and make predictions from the input RSI pairs. Moreover, AWDA integrates three essential cross-domain learning strategies to extract domain-invariant CD representations:

- **Supervised Learning.** It utilizes all labeled data from the source domain to train the model, aiming to establish its initial CD capability.
- **Domain Adversarial Training.** It uses an adversarial learning approach where a domain discriminator is trained to distinguish between source and target domain features, while the feature extractor is optimized to confuse the discriminator. This strategy enhances the extraction of domain-invariant features.
- **Class-weighted Self-Training.** To reduce prediction uncertainty of the unlabeled target domain, alongside addressing imbalanced class distribution, this strategy dynamically computes and assigns class weights during the self-training on the target domain, prioritizing the predictions of the minority class of change. Therefore, it promotes the class-balanced predictions in the target domain.

These three synergistic strategies enable AWDA to effectively transfer knowledge from a labeled source domain to an unlabeled target domain, mitigating domain-specific variations and enhancing the model's adaptability in diverse DACD scenarios.

To assess the effectiveness of the proposed AWDA framework in DACD, we established four cross-domain CD scenarios based on three well-known CD datasets: LEVIR, WHU, and GZ. These scenarios include LEVIR→WHU, WHU→LEVIR, LEVIR→GZ, and WHU→GZ. Utilizing these configurations, we conducted a thorough ablation study, hyper-parameter analysis, and comparative experiments.

To the best of our knowledge, our work is the first exploration on the cross-dataset CD task and we hope it is able to give some inspiration to future research, potentially broadening the application of CD in real-world scenarios.

The contributions are summarized as follows:

1. This study introduces a practical task of domain adaptive change detection (DACD), aiming to enable change knowledge transfer from a labeled source CD dataset to an unlabeled target CD dataset.
2. To reduce domain shifts and class imbalance during change knowledge from the source domain to the target domain, a novel Adversarial and Weighted Domain Adaptation (AWDA) framework is proposed for DACD.

3. The proposed AWDA significantly enhances CD performance on the target dataset without requiring its labels, achieving IoU improvements ranging from 13.64 to 34.73. It offers a clear advantage over competing DA methods across different DACD scenarios. This improvement notably boosts the practicality and feasibility of DACD.

## 2. Related work

### 2.1. Domain adaptation

DA has raised great attention in recent years as its powerful ability to reduce the need for time-consuming annotations in target dataset. Its primary objective is to address the discrepancy among different domains, thereby enhancing model's generalizability and transferability. In the current era of deep learning, numerous DA techniques have been developed, and they can be roughly divided into: (1) Explicit-alignment-based methods (Gretton et al., 2006; Huang et al., 2023a), which use some explainable strategies, such as feature alignment (Pan et al., 2010; Gretton et al., 2012a) and feature clustering (Tang et al., 2020), to align the cross-domain features for reducing domain shifts. These methods rely on manually designed subspace representations and explicit alignment in the latent space across domains. While they offer interpretability and control over the alignment process, their reliance on handcrafted subspaces can limit flexibility and scalability in complex scenarios; (2) Adversarial-training-based methods (Long et al., 2015; Ganin et al., 2016; Zhao et al., 2020; Chen et al., 2022; HassanPour Zonoozi and Seydi, 2023; Qu et al., 2024b), which aim to implicitly align source and target domain distributions using two networks: a generator that extracts features from both domains to confuse a discriminator, and a discriminator that attempts to distinguish between them. These methods provide an automated approach to cross-domain alignment and do not require labels for target domain. However, they often struggle with imbalanced class distributions, leading to suboptimal alignment for underrepresented classes; and (3) Self-training-based methods (Kumar et al., 2020; Huang et al., 2023a; Lee and Lee, 2023; Zhao et al., 2024), which optimize the network using pseudo-labels generated for the target domain to guide training. These methods effectively reduce prediction uncertainty and enhance performance on the target domain. Nevertheless, they generally assume that the labeled and unlabeled data share the same distribution, which can limit their applicability in scenarios with significant domain discrepancies. A detailed analysis of some certain methods used in the DACD experiments is provided in Section 4.8.

DA has been first employed in classification task (Gretton et al., 2006; Kang et al., 2019; Liang et al., 2021; Zhang et al., 2023c; Li et al., 2023; Zhou et al., 2024). For example, Maximum Mean Discrepancy (MMD) (Pan et al., 2010; Gretton et al., 2006, 2012a) is a method for measuring difference between two probability distributions. In transfer learning, minimizing MMD helps models adapt to different data distributions without requiring either source or target labels. A task-specific boundary alignment method was introduced to align the distributions between the source and target domains (Saito et al., 2018). A progressive feature alignment network (Chen et al., 2019) was designed to leverage cross-domain class consistency and address intra-class variance in the target domain. Domain consensus clustering was developed to enhance discriminative clustering in the target domain by utilizing sample-level and semantic-level domain consensus knowledge (Li et al., 2021). Recently, a source domain expansion strategy was presented to embed a robust prior by integrating pseudo-source data early in the adaptation process (Westfechtel et al., 2024).

Later, DA has been further performed on semantic segmentation tasks (Li et al., 2019; Guo et al., 2021; Xia et al., 2023; Zeng et al., 2024). A self-training-based DA architecture was presented (Zou et al., 2018), incorporating a class-balanced self-training strategy to address

the issue of imbalanced class semantic transfer. A multi-source adversarial domain aggregation network was proposed (Zhao et al., 2019) to facilitate adaptation from synthetic data to real-world data. A custom-designed network, DAFormer, was introduced (Hoyer et al., 2022) to explore the transfer capabilities of Transformers for semantic segmentation. An adaptive dual-path learning method was developed (Cheng et al., 2023), leveraging two mutually reinforcing paths to enhance adaptation ability. Additionally, a pseudo-label refinement network based on Fourier-based adaptation and contrastive learning was proposed (Zhao et al., 2024), enabling the online refinement of pseudo labels for DA in semantic segmentation.

## 2.2. Change detection

Over the past years, CD has garnered considerable attention owing to its substantial practical applications. As the dominant of this field, deep learning-based CD methods (Daudt et al., 2018; Zhang et al., 2020; Qu et al., 2023; Ning et al., 2024) have demonstrated encouraging performance. Fully-supervised learning methods typically require large amounts of annotated data for model training. For instance, Daudt et al. (2018) designed three fully convolutional neural network frameworks for CD trained from scratch in an end-to-end manner. Based on UNet (Ronneberger et al., 2015) architecture, Zheng et al. (2021) introduced Cross Layer Blocks (CLBs) for better integration of multi-level and multi-scale information. Chen et al. (2021) presented a transformer-based model to connect the semantic concepts across a Siamese-based spatiotemporal domain. Besides, to alleviate the heavy dependence on labels, semi-supervised change detection (SSCD) (Peng et al., 2020; Bandara and Patel, 2022; Yuan et al., 2024; Zuo et al., 2024; Huang et al., 2023b) utilizing both the scarce labeled data and the abundant unlabeled data for model training have also been developed. For example, Zhang et al. (2023b) proposed a model-free feature-prediction alignment (FPA) method for SSCD, of which prediction and feature alignments are designed to efficiently reduce the prediction uncertainty of unlabeled data. They further developed a joint self-training and rebalanced consistency learning (ST-RCL) approach to mitigate the adverse impact caused by the imbalanced distribution and rotation non-equivariance issue in SSCD. However, SSCD still needs a few labeled data to help the CD model obtain preliminary CD ability, limiting its rapid employment in unlabeled scenes.

Recently, there have been some attempts on reducing domain shifts between bi-temporal images of CD within the same dataset. For instance, some work attempted to reduce domain shifts caused by seasons (Kou et al., 2020; Huang and Zhang, 2024). A multi-source CD network was proposed in Zhang et al. (2022) to handle the diversity between SAR and optical images by using DA constraints to align diverse data into a unified deep feature space. A supervised DA architecture (Liu et al., 2022) was proposed for cross-domain CD within the same datasets, which employs a dual adaptation approach modifying feature and image attributes. A cycle-refined multi-decision joint alignment network was designed to mitigate the domain shifts in UDA hyperspectral CD (Qu et al., 2024a).

There are notable differences between the aforementioned SSCD, cross-domain CD, and the DACD approach proposed in this paper. SSCD focuses on leveraging unlabeled data that shares the same distribution as labeled data for model training, with both labeled and unlabeled data typically originating from the same dataset. Cross-domain CD, on the other hand, addresses the challenge of reducing domain shifts between pre-change and post-change images within the same dataset, where the training set is fully labeled. In contrast, our DACD approach seeks to transfer change detection knowledge from a fully-labeled source CD dataset to an unlabeled target CD dataset, enabling effective adaptation across datasets.

## 3. Adversarial and weighted domain adaptation for change detection

This section starts by defining the essential notations in DACD. Next, the encoder–decoder-based CD model is introduced. Then, three cross-domain learning strategies of AWDA are introduced in details, with a summary of the whole training procedure. The overall workflow of AWDA is shown in Fig. 2.

### 3.1. Notations

In the context of DACD, we define two distinct yet relevant domains (datasets): (1) a source domain, symbolized as  $D_s = \{(\mathbf{x}_a^s, \mathbf{x}_b^s, \mathbf{y}^s)\}_{n=1}^{N_s}$ , which includes both RSI pairs and their labels, and (2) a target domain, denoted as  $D_t = \{(\mathbf{x}_a^t, \mathbf{x}_b^t)\}_{n=1}^{N_t}$ , comprising solely RSI pairs without labels. In this notation,  $\mathbf{x}_a$  and  $\mathbf{x}_b$  represent a pair of RSIs, one taken before an event (pre-event image) and the other after (post-event image), while  $\mathbf{y}_i$  denotes the associated pixel-wise label map, with  $N$  indicating the number of samples in each dataset. The label map  $\mathbf{y}$  contains two values, [0,1], representing the classes of non-change and change, respectively.

For the CD model, our study employs an encoder–decoder-based Siamese network, consisting of three modules: (1) an encoder  $\mathcal{G}_x$  that extracts bi-temporal difference feature maps from RSI pairs, (2) a class decoder  $\mathcal{G}_y$  that projects these feature maps into pixel-wise prediction maps of change and non-change, and (3) a domain discriminator  $\mathcal{G}_d$  that projects the feature maps into either source or target domain.

### 3.2. Encoder-decoder CD architecture

As previously mentioned, both the source and target domains utilize the same encoder–decoder CD model, adhering to a uniform process for feature extraction and change prediction. A convolutional neural network (CNN)-based encoder is first used to derive feature maps from a bi-temporal RSI pair  $(\mathbf{x}_a, \mathbf{x}_b)$ . After it, the decoder enriched with a pyramid pooling module (PPM) as introduced by Zhao et al. (2017) and several upsampling modules as detailed by Shi et al. (2016) is employed to transform these feature maps into a change prediction map.

Firstly, the encoder  $\mathcal{G}_x$  is applied to a bi-temporal RSI pair, comprising a pre-event image  $\mathbf{x}_a \in \mathcal{R}^{H \times W \times 3}$  and a post-event image  $\mathbf{x}_b \in \mathcal{R}^{H \times W \times 3}$ , to extract two individual feature maps  $\mathbf{f}_a \in \mathcal{R}^{H/s \times W/s \times C}$  and  $\mathbf{f}_b \in \mathcal{R}^{H/s \times W/s \times C}$ . This process is mathematically represented as

$$\begin{cases} \mathbf{f}_a = \mathcal{G}_x(\mathbf{x}_a), \\ \mathbf{f}_b = \mathcal{G}_x(\mathbf{x}_b), \\ \mathbf{f}_d = |\mathbf{f}_a - \mathbf{f}_b|. \end{cases} \quad (1)$$

Here,  $H \times W$  denotes the spatial dimensions of the image,  $C$  represents the feature dimension, and  $s$  is the spatial shrinking ratio, determined by certain CD backbones. Absolute difference operation,  $|\mathbf{f}_a - \mathbf{f}_b|$ , is utilized to compute the temporal difference feature map  $\mathbf{f}_d$ , which can capture the essential variations between two temporal states.

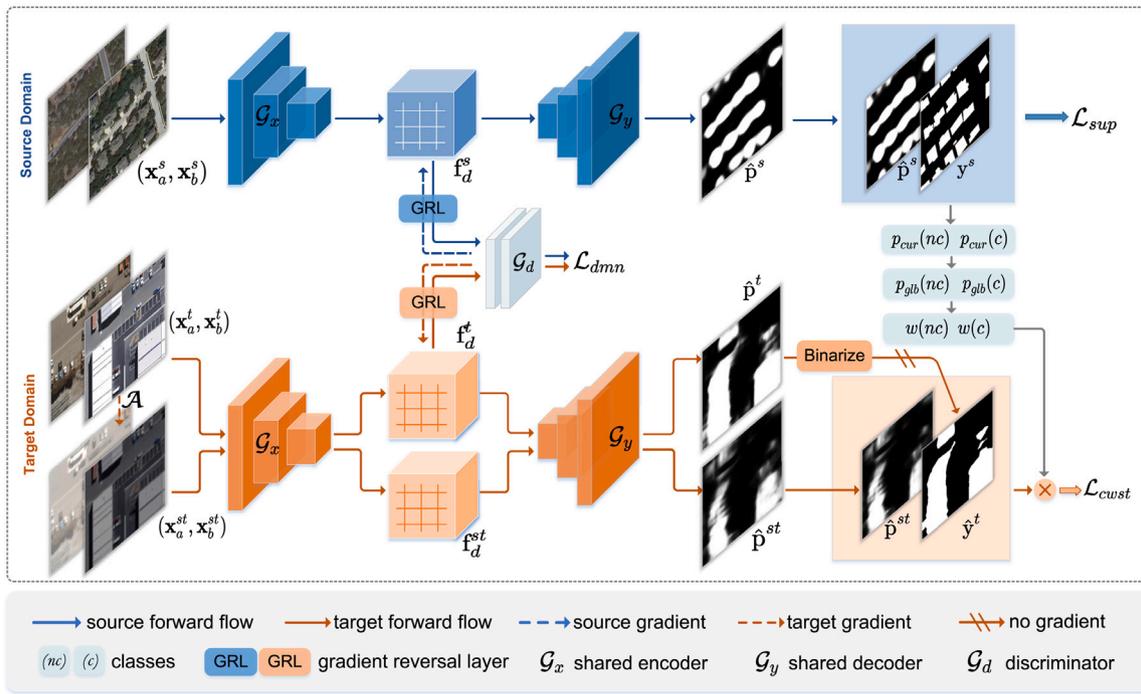
Then, the decoder  $\mathcal{G}_y$  decodes the difference feature map  $\mathbf{f}_d$  into a logit map of change activation  $\mathbf{p} \in \mathcal{R}^{H \times W \times 2}$ , where the dimension 2 represents the two classes of change and non-change. This operation can be formulated as

$$\mathbf{p} = \mathcal{G}_y(\mathbf{f}_d). \quad (2)$$

Finally, the change logit map  $\mathbf{p}$  undergoes the Softmax normalization to a probability map  $\hat{\mathbf{p}}$ . It ensures that the probabilities for the two classes of change and non-change at each location sum up to 1. The expression of this process is formulated as

$$\hat{p}_{ij} = \text{Softmax}(\mathbf{p}_{ij}) = \frac{e^{p_{ij}}}{\sum_{k=1}^2 e^{p_{ijk}}}, \quad (3)$$

where  $[i, j]$  is the spatial location index.



**Fig. 2.** The workflow AWDA for domain adaptive change detection. It consists of three key cross-domain learning strategies: (1) Supervised learning, optimized by  $\mathcal{L}_{sup}$ , utilizes all labeled data from the source domain to train the model and establish initial CD capability; (2) Domain adversarial training, optimized by  $\mathcal{L}_{dmn}$ , aligns cross-domain features adversarially to extract domain-invariant features; and (3) Class-weighted self-training, optimized by  $\mathcal{L}_{cwst}$ , dynamically computes and assigns class weights for the self-training on unlabeled data of the target domain to reduce their prediction uncertainty with respect to class imbalance.

### 3.3. Supervised learning of source domain

In supervised learning, the CD model is trained by all the training data of the source domain. For a source RSI pair  $(x_a^s, x_b^s)$ , its difference feature map  $f_d^s$  can be obtained by Eq. (1), and then its change probability map  $\hat{p}^s$  is generated by Eqs. (2)–(3). The probability map  $\hat{p}^s$  is utilized to optimize the CD model through Cross-Entropy (CE) loss, which is mathematically formulated as

$$\mathcal{L}_{sup}(G_x, G_y) = \frac{1}{HW} \sum_{i=1}^H \sum_{j=1}^W \text{CE}[y^s(i, j), \hat{p}^s(i, j)]. \quad (4)$$

It ensures that the CD model,  $G_x$ - $G_y$ , learns from the labeled data of the source domain and acquires a preliminary CD ability. However, due to the domain shifts, the CD model cannot be directly adapted to the target domain without DACD.

### 3.4. Adversarial learning between source and target domains

In Section 3.2, each source RSI pair  $(x_a^s, x_b^s)$  and each target RSI pair  $(x_a^t, x_b^t)$  can generate a source difference feature map  $f_d^s$  and a target difference feature map  $f_d^t$  by Eq. (1), respectively. There are domain shifts between  $f_d^s$  and  $f_d^t$ , hindering model generalization across datasets. To reduce domain shifts at the feature level, This subsection introduces adversarial learning between  $f_d^s$  and  $f_d^t$  based the strategy of Gradient Reversal Layer (GRL) (Ganin et al., 2016).

First, the domain discriminator  $G_d$  classify whether the extracted difference feature map  $f_d$  originates from the source or target domain. Given  $f_d^s$  and  $f_d^t$ ,  $G_d$  outputs the corresponding domain probability distributions  $G_d(f_d^s)$  and  $G_d(f_d^t)$ , respectively. Then, a domain adversarial loss  $\mathcal{L}_{dmn}$  is calculated and minimized to encourage the discriminator  $G_d$  to correctly classify domain origins. The loss is formulated as

$$\mathcal{L}_{dmn}(G_x, G_d) = -\mathbb{E}_{f_d^s}[\log G_d(f_d^s)] - \mathbb{E}_{f_d^t}[\log(1 - G_d(f_d^t))]. \quad (5)$$

However, directly using  $\mathcal{L}_{dmn}$  to optimize the encoder  $G_x$  shared between source and target domains would unexpectedly enlarge the

domain shifts between  $f_d^s$  and  $f_d^t$ . To reversely promote  $G_x$  to extract domain-invariant features, the GRL is used to reverse the gradients propagated during their backpropagation between  $G_d$  to  $G_x$  as

$$\frac{\partial \mathcal{L}_{dmn}}{\partial f_d^s}, \frac{\partial \mathcal{L}_{dmn}}{\partial f_d^t} = -\lambda \cdot \frac{\partial \mathcal{L}_{dmn}}{\partial G_d(f_d^s)}, -\lambda \cdot \frac{\partial \mathcal{L}_{dmn}}{\partial G_d(f_d^t)}. \quad (6)$$

Here,  $\mathcal{L}_{dmn}$  represents the above domain adversarial loss.  $\frac{\partial \mathcal{L}_{dmn}}{\partial G_d(f_d^s)}$  and  $\frac{\partial \mathcal{L}_{dmn}}{\partial G_d(f_d^t)}$  are the original gradients of the loss with respect to  $G_d(f_d^s)$  and  $G_d(f_d^t)$ . The term  $\lambda$  is a scalar controlling the strength of gradient reversal, following a ramp-up strategy from 0 to 1 based on the ratio of current iterations to total iterations as

$$\lambda = \frac{2}{1 + e^{-10 \cdot \frac{\text{curr\_iter}}{\text{total\_iters}}}} - 1. \quad (7)$$

As a result, the GRL-based domain adversarial training effectively redistributes gradients to promote domain-invariant feature extraction by  $G_x$ . By reversing the gradients of domain discrimination,  $G_x$  is encouraged to update its parameters in a manner that minimizes domain-specific differences captured by  $G_d$ . This process facilitates the learning of domain-invariant features, improving the model's generalization capability.

### 3.5. Class-weighted self-training of target domain

To further reduce the prediction uncertainty of the CD model on the target domain, it is helpful to employ self-training) on the target domain. As discussed in Section 1, imbalanced distribution between change and non-change could cause the model to exhibit significant bias towards the majority class of non-change during training. To mitigate this issue, we develop a novel approach called Class-weighted Self-training (CWST). It integrates an adaptive class weight strategy, which dynamically computes class-specific weights for non-change and change based on their convergence difficulty, into the self-training framework.

### 3.5.1. Adaptive class weight calculation

The adaptive class calculation strategy consists of the following three steps: (1) computing the average class probabilities of the source domain within each mini-batch, (2) employing Exponential Moving Average (EMA) to continuously update these batch class average probabilities towards the corresponding global class average probabilities, and (3) translating the global class average probabilities to corresponding class weights.

**Batch Average Class Probability.** During each mini-batch of the source domain with the batch size set to  $B$ , there exists a batch comprising prediction probabilities  $\hat{\mathbf{P}} = [\hat{\mathbf{p}}_1, \hat{\mathbf{p}}_2, \dots, \hat{\mathbf{p}}_B]$  and corresponding labels  $\mathbf{Y} = [y_1, y_2, \dots, y_B]$ , and it can generate class average probabilities  $p_{cur} \in \mathcal{R}^2$ . The average class probability within this mini-batch can use the following formulation:

$$p_{cur}(c) = \frac{\sum_{b=1}^B \sum_{i=1}^H \sum_{j=1}^W \hat{\mathbf{P}}(b, i, j) \cdot \mathbb{1}[\mathbf{Y}(b, i, j) = c]}{\sum_{b=1}^B \sum_{i=1}^H \sum_{j=1}^W \mathbb{1}(\mathbf{Y}(b, i, j) = c)}, \quad (8)$$

where  $c$  represents the class index, which is either 0 or 1. The expression  $\mathbb{1}[\mathbf{Y}(b, i, j) = c]$  constitutes an indicator function, the value of which is 1 when  $\mathbf{Y}(b, i, j) = c$ , and 0 otherwise. It can calculate the class-wise average probability value in  $\hat{\mathbf{P}}$  for all the pixels where  $y$  equals class- $c$ .

**Global Average Class Probability.** In order to enhance the stability of the representation of class probabilities, the batch average class probabilities  $p_{cur}$  are updated to their corresponding global average class probabilities, denoted as  $p_{glob}$ , which are initialized with 1, as

$$p_{glob}(c) = \alpha p_{glob}(c) + (1 - \alpha) p_{cur}(c), \quad (9)$$

Where  $\alpha$  is the updating momentum, the value of which is set to 0.99. It is worthy of note that class- $c$  shall not be updated when the class does not exist during the current iteration.

**Class Weight.** Based on the stable global average class probabilities  $p_{glob}$ , we can translate them to corresponding class weight denoted as  $\mathbf{w} \in \mathcal{R}^2$  at difference stages during training. For class- $c$ , its weight  $\mathbf{w}(c)$  is calculated by

$$\mathbf{w}(c) = \frac{1}{p_{glob}(c)^{2(1 - \frac{curr\_iter}{total\_iters}) + 1}}, \quad (10)$$

where the weight  $\mathbf{w}(c)$  decreases as  $p_{glob}(c)$  increases. In other words, as the CD model's confidence in the predictions of class- $c$ ,  $p_{glob}(c)$ , grows stronger, the assigned weight for the class,  $\mathbf{w}(c)$ , decreases correspondingly. Besides,  $\mathbf{w}(c)$  is a monotonically decreasing function with training iterations. When the current iteration approaches the total iterations, there exists approximately  $\mathbf{w}(c) \rightarrow \frac{1}{p_{glob}(c)}$ .

### 3.5.2. Class-weighted self-training

After obtaining the class weight  $\mathbf{w}$ , we can assign it to the self-training of the unlabeled target domain. Inspired by the workflow of FixMatch (Sohn et al., 2020), a classical paradigm in semi-supervised learning, we use two spatial-aligned for robust self-training. The two branches include a weakly-augmented target RSI pair  $(\mathbf{x}_a^t, \mathbf{x}_b^t)$  and a strongly-augmented target RSI pair  $(\mathbf{x}_a^{st}, \mathbf{x}_b^{st})$ .

In this study, the weakly-augmented RSI pair  $(\mathbf{x}_a^t, \mathbf{x}_b^t)$  equals that introduced in Section 3.4. Its probability prediction map, denoted as  $\hat{\mathbf{p}}^t$ , can be obtained using Eqs. (1)–(3). Subsequently, its pseudo-label map  $\hat{\mathbf{y}}^t \in \mathcal{R}^{H \times W}$  at the location  $[i, j]$  is generated as

$$\hat{\mathbf{y}}^t(i, j) = \arg \max_{c \in \{0,1\}} \hat{\mathbf{p}}^t(i, j, c), \quad (11)$$

The strongly-augmented RSI pair  $(\mathbf{x}_a^{st}, \mathbf{x}_b^{st})$  is generated from  $(\mathbf{x}_a^t, \mathbf{x}_b^t)$  by some random strong augmentations represented as  $\mathcal{A}$ , which is formulated as

$$\mathbf{x}_a^{st}, \mathbf{x}_b^{st} = \mathcal{A}(\mathbf{x}_a^t), \mathcal{A}(\mathbf{x}_b^t), \quad (12)$$

where  $\mathcal{A}$  is two connected intensity augmentations randomly selected from a strong augmentation list (Cubuk et al., 2020), following the

setting of FPA (Zhang et al., 2023b). The probability prediction map of  $(\mathbf{x}_a^{st}, \mathbf{x}_b^{st})$ , denoted as  $\hat{\mathbf{p}}^{st}$ , can be obtained via Eqs. (1)–(3) as that of  $(\mathbf{x}_a^t, \mathbf{x}_b^t)$ .

For a target RSI pair, after obtaining the probability maps of weakly- and strongly-augmented versions,  $\hat{\mathbf{p}}^t$  and  $\hat{\mathbf{p}}^{st}$ , along with its pseudo-label map  $\hat{\mathbf{y}}^t$ , the class-weighted self-training loss is calculated as

$$\mathcal{L}_{cwtst}(G_x, G_y) = \frac{1}{HW} \sum_{i=1}^H \sum_{j=1}^W \mathbf{w}[\hat{\mathbf{y}}^t(i, j)] \cdot \mathbb{1}[\hat{\mathbf{p}}^t(i, j) > \tau] \cdot \text{CE}[\hat{\mathbf{y}}^t(i, j), \hat{\mathbf{p}}^{st}(i, j)]. \quad (13)$$

Here, we utilize the pseudo-labels from weakly-supervised RSI pairs to supervise the predictions of strongly-augmented RSI pairs. To mitigate the influence of noisy pseudo-labels, only those with high confidence are selected, determined by the indicator function  $\mathbb{1}[\hat{\mathbf{p}}^t(i, j) > \tau]$ . Additionally, the weight  $\mathbf{w}[\hat{\mathbf{y}}^t(i, j)]$  of the pseudo-label class at position  $(i, j)$  is assigned to guide the self-training process at this point. When the term  $\mathbf{w}$  is absent, the loss simplifies to the normal self-training loss of FixMatch, denoted as  $\mathcal{L}_{st}$ .

### 3.6. Overall loss and training procedure

The overall loss of the proposed AWDA is the sum of the supervised loss of source domain  $\mathcal{L}_{sup}$ , the domain adversarial loss  $\mathcal{L}_{dmn}$ , and the class-weighted self-training loss  $\mathcal{L}_{cwtst}$ , as

$$\mathcal{L} = \mathcal{L}_{sup} + \mathcal{L}_{dmn} + \mathcal{L}_{cwtst}. \quad (14)$$

The whole training procedure of the proposed AWDA for DACD is summarized in Algorithm 1.

#### Algorithm 1 Training procedure of AWDA

**Input:** source domain  $D_s$ , target domain  $D_t$ , encoder-decoder-based CD model  $G_x-G_y$ , domain discriminator  $G_d$ , epoch number  $N_{epoch}$ , iteration number of each epoch  $N_{iter}$

```

for epoch  $\leftarrow$  1 to  $N_{epoch}$  do
  for iteration  $\leftarrow$  1 to  $N_{iter}$  do
    Data Processing: sample source RSI pairs, sample and
    transform target RSI pairs via Eq. (12);
    Supervised Learning of Source Domain: calculate the super-
    vised loss  $\mathcal{L}_{sup}$  between source RSI pairs and their labels via
    Eqs. (1)-(4);
    Cross-domain Adversarial Learning: calculate the domain ad-
    versarial loss  $\mathcal{L}_{dmn}$  between source RSI pairs and target RSI
    pairs via Eqs. (1) and (5)-(7);
    Class-weighted Self-training of Target Domain:
      (I) Adaptive Class Weight Calculation: calculate class weights
      via Eqs. (8)-(10);
      (II) Class-weighted Self-training: assign the class weights to
      the self-training on the target domain and therefore calculate
      class-weighted self-training loss  $\mathcal{L}_{cwtst}$  via Eqs. (11)-(13);
    Model Optimization: calculate the overall loss  $\mathcal{L}_{sup} + \mathcal{L}_{dmn} +$ 
     $\mathcal{L}_{cwtst}$  via Eq. (14) and use it to optimize  $G_x-G_y-G_d$ .
  end
end
Output: optimized CD model  $G_x-G_y$ 

```

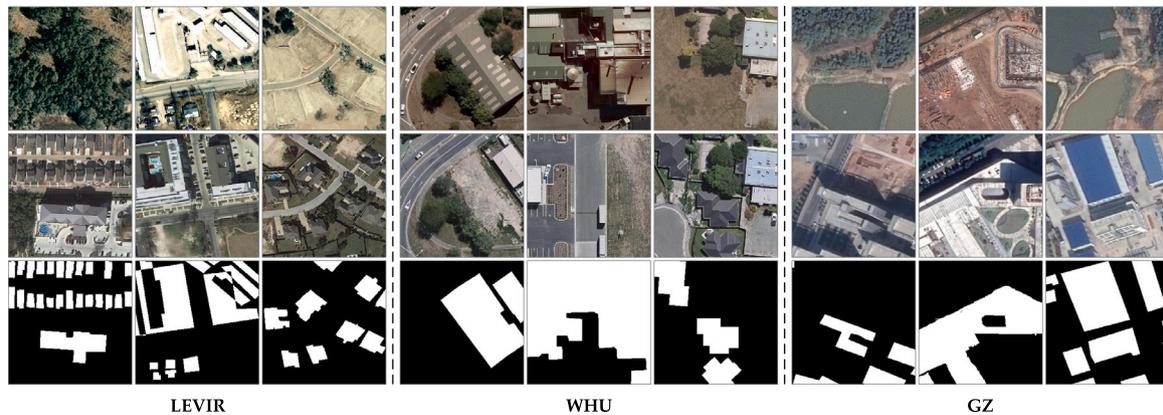
## 4. Experiments

### 4.1. Datasets

Three building CD datasets, WHU (Ji et al., 2018), GZ (Peng et al., 2020), and LEVIR (Chen and Shi, 2020), are utilized for constructing cross-domain CD scenarios in experiments to verify the cross-domain CD performance of the model. Their characteristics and training/validation/test splits are summarized in Table 1. Each dataset is randomly

**Table 1**  
Detailed characteristics of three used datasets.

Dataset	LEVIR	WHU	GZ
Image mode	RGB	RGB	RGB
Capturing time	2002–2018	2012–2016	2006–2019
Included area	Texas (America)	Christchurch (New Zealand)	Guangzhou (China)
Pixel ratio of change	4.67%	4.33%	8.5%
Spatial resolution	0.5 m/pixel	0.075 m/pixel	0.55 m/pixel
Original size	1024 × 1024	32507 × 15354	1006 × 1168 - 4936 × 5224
Original image pairs	637	1	19
Cropped size	256 × 256	256 × 256	256 × 256
Cropped image pairs	10192	7437	3603
Ratio of train:Validation:Test	6:2:2	6:2:2	6:2:2
Pairs of train:Validation:Test	6115:2038:2039	4460:1486:1488	2161:720:722



**Fig. 3.** Examples from the LEVIR, WHU, and GZ datasets are displayed. The pre-temporal images, the post-temporal images, and the corresponding ground truths are in the first row, the second row, and the third row, respectively.

split into training, validation, and test sets without overlapping. From the perspective of domain shifts, it can be found that the datasets exhibit significant variations, ranging from spatial resolution differences (from 0.55 m/pixel in GZ to 0.075 m/pixel in WHU) to regional differences (from Texas to Guangzhou). These disparities underscore the diversity in change semantics across different datasets.

Based on the three datasets, we constructed four cross-domain scenarios for DACD, including: LEVIR→WHU, LEVIR→GZ, WHU→LEVIR, and WHU→GZ, where the former serves as source domain and the latter serves as target domain. In each scenario, the labeled training set from the source domain and the unlabeled training set from the target domain are utilized for model training. Meanwhile, the validation and test sets of the target domain are employed for model selection and evaluation (see Fig. 3).

#### 4.2. Evaluation metrics

Five commonly-used CD evaluation metrics are employed to assess the performance of all the methods in the experiments: Recall, Precision, Intersection over Union (IoU), F1-Score (F1), and the Kappa coefficient (Kappa). The IoU, F1, Recall, and Precision metrics are measured on a scale from 0 to 100%, while the Kappa coefficient ranges between -1 and 1. For all these metrics, a higher value indicates superior CD performance. They are formulated as

$$Recall = \frac{TP}{FN + TP}, \tag{15}$$

$$Precision = \frac{TP}{FP + TP}, \tag{16}$$

$$IoU = \frac{Recall \times Precision}{Recall + Precision - Recall \times Precision}, \tag{17}$$

$$F1 = \frac{2 \times Recall \times Precision}{Recall + Precision}, \tag{18}$$

$$OA = \frac{TP + TN}{FP + FN + TP + TN}, \tag{19}$$

$$PRE = \frac{(TN + FN) \times (FP + TN)}{(TN + TP + FN + FP)^2} + \frac{(TP + FP) \times (FN + TP)}{(TN + TP + FN + FP)^2}, \tag{20}$$

$$Kappa = \frac{OA - PRE}{1 - PRE}, \tag{21}$$

where  $TN$  and  $TP$  denote the accurately identified counts of unchanged and changed pixels, respectively. On the other hand,  $FN$  represents the count of changed pixels that are erroneously classified as unchanged, and  $FP$  corresponds to the total number of unchanged pixels incorrectly identified as changed. This study focuses on the area of change, therefore, only the metrics of the change class are calculated and reported.

#### 4.3. Experimental settings

In comparison experiments, all the competing methods including our method use Dilated ResNet50 (He et al., 2016) pre-trained on ImageNet (Deng et al., 2009) as the Siamese encoder to extract features with the spatial shrinking ratio  $s$  set to 4. The architecture of all the decoders, including the domain decoder, class decoder, and domain-class decoder, is identical, with the only difference being the number of outputs in the final layer. The detailed architecture of each decoder comprises the following sequence: a  $3 \times 3$  convolution layer followed by a ReLU activation, an upsample layer, and a final  $1 \times 1$  convolution layer. Throughout, the middle feature dimensions are uniformly set to 32. The threshold  $\tau$  in Eq. (13) is set to 0.95 for all the adaptation scenarios.

For objective comparison, the encoder–decoder model is consistently trained across all methods for a total of 50 epochs, using a mini-batch size of 8; the discriminators of the comparison methods

Table 2

Ablation study of the proposed AWDA with ResNet50 as the encoder backbone in two cross-domain CD scenarios.

Source→Target				LEVIR→WHU					LEVIR→GZ				
$\mathcal{L}_{sup}$	$\mathcal{L}_{dmn}$	$\mathcal{L}_{st}$	$\mathcal{L}_{cwst}$	IoU	F1	Kappa	Rec.	Pre.	IoU	F1	Kappa	Rec.	Pre.
✓				49.64	66.35	0.6492	58.07	77.38	27.52	43.16	0.4055	29.14	83.16
✓	✓			50.92	67.48	0.6585	67.75	67.21	36.79	53.80	0.5109	40.00	82.11
✓		✓		11.08	19.96	0.1898	11.43	78.75	49.33	66.07	0.6333	59.22	74.71
✓	✓	✓		59.73	74.79	0.7362	69.72	80.64	58.38	73.72	0.7156	66.91	82.07
✓	✓		✓	<b>70.73</b>	<b>82.85</b>	<b>0.8202</b>	<b>80.51</b>	<b>85.34</b>	<b>62.25</b>	<b>76.73</b>	<b>0.7483</b>	<b>69.24</b>	<b>86.03</b>

follow their original literature with the same learning rates of adversarial learning. The training on the encoder–decoder-based CD model employs Stochastic Gradient Descent (SGD) as the optimizer, with an initial learning rate of 0.01, which decreases according to a weight decay of  $1 \times 10^{-4}$  and a momentum of 0.9. All experiments are conducted using PyTorch 1.9.0 and are performed on two GeForce RTX 2080Ti GPUs.

#### 4.4. Ablation study and effects of AWDA

To verify the effectiveness of each component of the proposed AWDA, we conducted its ablation studies with ResNet50 as its encoder backbone. As introduced in Section 3.6, there are three losses in AWDA, including the supervised loss of source domain  $\mathcal{L}_{sup}$  (SUP module), domain adversarial loss  $\mathcal{L}_{dmn}$  (DMN module), class-weighted self-training loss of target domain  $\mathcal{L}_{cwst}$  (CWST module). To further study the effect of adaptive class weighting, the self-training loss without class weight in Eq. (13), i.e.,  $\mathcal{L}_{st}$ , is compared with  $\mathcal{L}_{cwst}$ . Ablation studies are carried out on two DACD scenarios of LEVIR→WHU and WHU→GZ, with experimental results provided in Table 2.

It was observed that compared to the base method of SUP, neither the individual DMN nor the ST/CWST contributes to stable and effective performance improvement. Simply applying AT to ST, i.e., AS+ST, makes a limited improvement of IoU in LEVIR→WHU from 49.64 to 50.92. What is more, directly adding the naive ST module to SUP, SUP+ST, even results in a significant decline in performance from 49.64 to 11.08 in LEVIR→WHU. These results reveal that individual feature alignment or self-training cannot achieve effective change knowledge transfer in complicated cross-domain scenarios, where both domain shifts and class imbalance exist. When implementing the SUP+DMN+ST configuration, there is a noticeable performance improvement. For example, its IoU increases significantly by 10.09, from 49.64 to 59.73 in the LEVIR→WHU scenario. Such results verify the importance of jointly reducing domain shifts and alleviating class imbalance. Finally, with the employment of CWST, the complete AWDA framework, SUP+DMN+CWST, achieves the stably optimal performances in both cross-domain scenarios. This highlights the efficacy of cross-domain feature alignment and class rebalancing.

To further explore how the proposed AWDA improves cross-domain CD performance, based on the cross-domain CD scenario LEVIR→WHU, we examine its effects from two perspectives that align with our motivations:

(1) **Reducing domain shifts.** As depicted in Fig. 4, we analyze the cosine similarity of the mean feature embeddings between the source and target domains. The results illustrate that the DMN module effectively mitigates domain shifts, leading to improved cross-domain feature similarity. However, the ST module noticeably decreases the similarity from 0.8946 in SUP to 0.8777 in SUP+ST. Interestingly, when DMN and ST are combined (SUP+DMN+ST), the similarity is restored to 0.8958. Finally, the integration of all modules (SUP+DMN+ST) achieves the highest similarity of 0.9208. These findings underscore the efficacy of AWDA's sub-modules in minimizing cross-domain shifts.

(2) **Alleviating class imbalance.** Self-training on the target domain is a prevalent strategy within DA methods, ideally aimed at reducing the uncertainty of predictions. However, the introduction of the ST module dramatically degrades CD performance in the target domain.

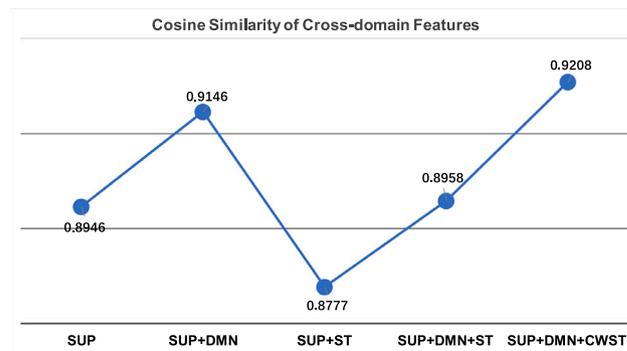


Fig. 4. Cosine similarity of the mean change feature values between source and target domains across various AWDA variants in the cross-domain CD scenario LEVIR→WHU. Higher similarity values indicate smaller domain shifts.

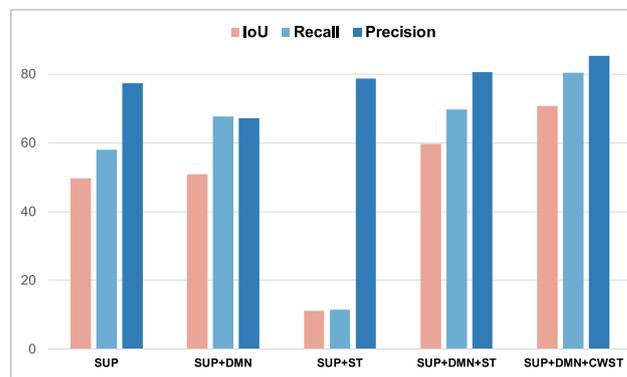


Fig. 5. IoU value varies with Recall and Precision in LEVIR→WHU for different components of AWDA.

To investigate the potential reasons for this, in Fig. 5, we plot the IoU along with its contributing factors, Recall and Precision, as described in Eq. (17). The figure demonstrates that the overall model performance, i.e., IoU, is constrained by Recall while Precision remains stably high. The low Recall value of SUP+ST suggests that the CD model exhibits a bias towards non-change, demonstrating significant class imbalance when relying solely on self-training. Fortunately, the integration of CWST effectively reverses this phenomenon, enhancing Recall to a similar level at Precision. These results further demonstrate the effect of the whole AWDA in alleviating class imbalance in DACD.

#### 4.5. Class weight of AWDA

In the proposed AWDA, class weight  $w$  that varies with the corresponding global class probability  $p$  during training is essential for addressing class imbalance in DACD. To intuitively demonstrate the adjustment effect, the weights for non-change and change classes are plotted against their class probabilities in Fig. 6.

It can be observed that the weight of the non-change class is slightly over 1 due to its high probability being close to 1. In contrast, the

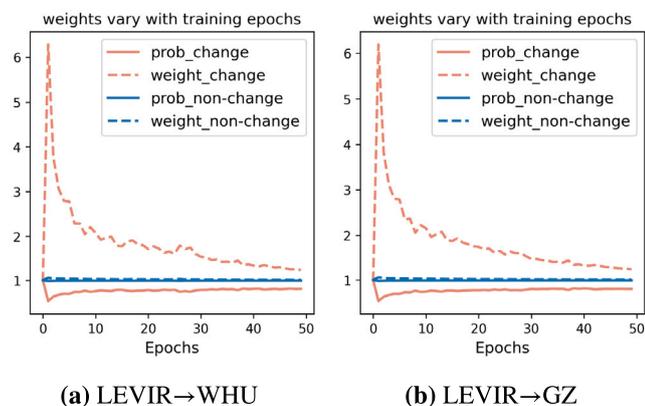


Fig. 6. Weights of non-change and change calculated in Eq. (9) vary with their global probabilities calculated in Eq. (10) at different training epochs in two DACD scenarios.

weight of the change class rapidly climbs to a peak value of more than 6 during the first few epochs, driven by its low probability and low confidence. Subsequently, this weight gradually decreases as the class probability increases. These results demonstrate that the proposed class-wise weighting strategy effectively mitigates class imbalance, as evidenced by the reduced class probability gap between non-change and change.

#### 4.6. Effects of different class weighting strategies

One of the key motivations of this paper is to address class imbalance, which negatively impacts the performance of change detection when transferring knowledge from the source domain to the target domain. To further investigate the effects of class imbalance, this study evaluates two additional class weighting strategies in the source domain: (1) **Fixed Weight**. This approach assigns constant weights to the non-change and change classes, with a higher weight allocated to the change class compared to the non-change class. Their weight sum is equal to 1. (2) **Decaying Weight**. This strategy applies a gradually decreasing weight to the change class, starting from 0.95 and reducing to 0.5 according to the formula  $w_c = 0.95 - 0.45 \left( \frac{\text{curr\_iter}}{\text{total\_iters}} \right)^2$ . The complementary weight is assigned to the non-change class, calculated as  $w_{nc} = 1 - w_c$ . These strategies aim to provide a deeper understanding of how class imbalance influences the model's performance.

Their results on LEVIR→WHU and LEVIR→GZ are reported in Table 3, in comparison with our AWDA. With a weight ratio of 0.5:0.5, which corresponds to the baseline method, the CD performance in the target domain deteriorates significantly, characterized by a much lower recall compared to the corresponding precision. For instance, in the LEVIR→GZ scenario, the recall is only 29.14, while the precision reaches 83.16. Assigning a higher weight to the change class effectively reduces the gap between recall and precision, leading to improved performance metrics such as IoU, F1-score, and Kappa. Notably, the best performance in the target domain is observed with weight ratios in the range of 0.1:0.9 to 0.2:0.8 for fixed weight. Both the fixed weight and decaying weight strategies effectively improve change detection performance in the target domain; however, the fixed weight strategy requires more precise tuning of hyperparameters to achieve optimal results. Compared to the two weighting methods applied in the source domain, our AWDA, which performs the weighting operation in the target domain, achieves higher recall and precision. This demonstrates the importance of further self-training in the target domain.

Overall, the comparison of different weighting strategies highlights the necessity of class weighting for DACD and underscores the advantages of applying further weighting training in the target domain.

#### 4.7. Comparison experiments

To objectively assess the effectiveness of the proposed AWDA method, we compared it against eight other methods, including one baseline and seven advanced DA techniques across four DACD scenarios: LEVIR→WHU, LEVIR→GZ, WHU→LEVIR, and WHU→GZ. To ensure a fair comparison, all the experiments were conducted using the same ResNet50 encoder backbone by default, unless otherwise specified. The comparison methods are:

- **OnlySrc**. This method utilizes the entire training dataset exclusively from the source domain for supervised training. It serves as the baseline for comparison in all subsequent methods, Its results are based on DeepLabV3+ with ResNet50 as the backbone.
- **OnlySrc-HRNet & OnlySrc-SegFormer**. Both methods adopt the same training settings as OnlySrc but utilize different backbones. Specifically, OnlySrc-HRNet employs HRNet-V2 (Wang et al., 2020), which has been previously applied to CD (Zhang et al., 2023b), while OnlySrc-SegFormer utilizes SegFormer-B1 (Xie et al., 2021).
- **FixMatch** (Sohn et al., 2020). It is a classical semi-supervised learning paradigm, and it is compared in this study to investigate the effect of cross-domain self-training on the target domain.
- **MMD** (Gretton et al., 2012b). It measures differences between probability distributions. Minimizing MMD helps models adapt to different data distributions without requiring labeled information.
- **DANN** (Ganin et al., 2016). It uses gradient reversal layers to achieve domain adversarial learning between source and target domains.
- **CLAN** (Luo et al., 2019), abbreviated as category-level adversarial network. It prioritize category-level alignment during domain adversarial training.
- **STADA** (Liang et al., 2023), abbreviated as self-training adversarial DA. It combines naive self-training and domain adversarial training for RS cross-domain semantic segmentation.
- **AdaptSegNet** (Tsai et al., 2018). It conducts domain adversarial training in the output space to directly align the predicted label distributions between source and target domains.
- **MemoryAdaptSegNet** (Zhu et al., 2023). Based on AdaptSegNet, it embeds an invariant feature memory module to store domain-level context information, which is integrated with a category attention-driven module to augment pixel representations.

Comparison results are shown in Tables 4 and 5. In general, the proposed AWDA surpasses all the other comparison approaches across all four adaptation scenarios. For instance, it attains an IoU score of 70.73 in LEVIR→WHU, thereby achieving a gain of 18.1 compared with the suboptimal method MemoryAdaptSegNet. In WHU→GZ, it achieves an IoU value of 50.43 which has an obvious advantage of 13.02 over the second-best method FixMatch. In LEVIR→GZ, it makes an enhancement in Kappa coefficient from 0.6333 to 0.7483 compared with the suboptimal approach FixMatch. It should be emphasized that our AWDA shows an improvement ranging from at least 13.64 to 34.74 in IoU compared to the OnlySrc method, which indicates the definite superiority of our approach.

Although FixMatch and DANN occasionally demonstrate notable performance, their results are generally unstable. For instance, FixMatch achieves an IoU score of 49.33 on the LEVIR→GZ adaptation task, outperforming the OnlySrc method by 21.81 points. However, its performance drops significantly in LEVIR→WHU, where it only attains an IoU score of 11.09, trailing OnlySrc by 38.55 points. In contrast, our proposed AWDA method consistently outperforms all baselines across the four adaptation scenarios, demonstrating its superior robustness.

Some visual examples of optimization-related methods (based on ResNet50-DeepLabV3+, by default) are provided in Fig. 7. Compared with other methods, our AWDA demonstrates consistent and reliable

**Table 3**  
Performance comparison of different class weighting strategies for OnlySrc in domain-adaptive change detection experiments.

Source→Target		LEVIR→WHU					LEVIR→GZ				
		IoU	F1	Kappa	Rec.	Pre.	IoU	F1	Kappa	Rec.	Pre.
Fixed weight	0.05:0.95	52.96	69.25	0.6769	70.29	68.23	40.95	58.11	0.5460	53.59	63.47
	0.1:0.9	54.55	70.59	0.6904	74.57	67.02	49.15	65.90	0.6278	65.42	66.39
	0.2:0.8	55.19	71.12	0.6964	73.15	69.20	46.53	63.51	0.6072	54.82	75.47
	0.3:0.7	51.35	67.86	0.6639	62.67	73.98	36.74	53.73	0.5097	40.29	80.65
	0.5:0.5 (baseline)	49.64	66.35	0.6492	58.07	77.38	27.52	43.16	0.4055	29.14	83.16
Decaying weight	0.95→0.5	55.87	71.69	0.7036	68.09	75.70	43.60	60.72	0.5694	63.13	58.49
Our AWDA		<b>70.73</b>	<b>82.85</b>	<b>0.8202</b>	<b>80.51</b>	<b>85.34</b>	<b>62.25</b>	<b>76.73</b>	<b>0.7483</b>	<b>69.24</b>	<b>86.03</b>

**Table 4**  
Comparison results when LEVIR works as the source domain with ResNet50 as the encoder backbone.

Source→Target		LEVIR→WHU					LEVIR→GZ				
		IoU	F1	Kappa	Rec.	Prec.	IoU	F1	Kappa	Rec.	Prec.
OnlySrc		49.64	66.35	0.6492	58.07	77.38	27.52	43.16	0.4055	29.14	83.16
OnlySrc-HRNet (Zhang et al., 2023b)		53.44	69.65	0.6823	66.03	73.70	33.75	50.47	0.4772	36.62	81.16
OnlySrc-SegFormer (Xie et al., 2021)		53.55	69.75	0.6842	62.70	78.58	37.34	54.38	0.5156	41.41	79.15
MMD (Gretton et al., 2012b)		47.61	64.51	0.6302	56.06	75.96	21.57	35.48	0.3317	22.30	<b>86.73</b>
FixMatch (Sohn et al., 2020)		11.09	19.96	0.1898	11.43	78.75	49.33	66.07	0.6333	59.22	74.71
DANN (Ganin et al., 2016)		50.92	67.48	0.6585	67.75	67.21	36.79	53.80	0.5109	40.00	82.11
CLAN (Luo et al., 2019)		45.60	62.63	0.6127	49.68	84.71	24.00	38.71	0.3617	25.34	81.94
STADA (Liang et al., 2023)		4.63	8.86	0.0822	4.78	60.53	12.95	22.94	0.2084	13.61	73.02
AdaptSegNet (Tsai et al., 2018)		48.31	65.14	0.6368	56.43	77.04	29.44	45.49	0.4296	30.85	86.55
MemoryAdaptSegNet (Zhu et al., 2023)		52.63	68.97	0.6771	58.54	83.92	31.47	47.87	0.4523	33.49	83.88
Our AWDA		<b>70.73</b>	<b>82.85</b>	<b>0.8202</b>	<b>80.51</b>	<b>85.34</b>	<b>62.25</b>	<b>76.73</b>	<b>0.7483</b>	<b>69.24</b>	<b>86.03</b>

**Table 5**  
Comparison results when WHU works as the source domain with ResNet50 as the encoder backbone.

Source→Target		WHU→LEVIR					WHU→GZ				
		IoU	F1	Kappa	Rec.	Prec.	IoU	F1	Kappa	Rec.	Prec.
OnlySrc		25.86	41.09	0.3811	40.85	41.34	36.79	53.79	0.5024	46.59	63.62
OnlySrc-HRNet (Zhang et al., 2023b)		20.46	33.98	0.3086	31.93	36.30	32.55	49.11	0.4448	48.44	49.80
OnlySrc-SegFormer (Xie et al., 2021)		15.22	26.42	0.2471	17.30	55.87	36.81	53.81	0.5077	42.39	73.64
MMD (Gretton et al., 2012b)		24.50	39.36	0.3666	35.12	44.77	31.78	48.23	0.4442	40.61	59.39
FixMatch (Sohn et al., 2020)		18.20	30.80	0.2884	21.52	54.19	37.41	54.45	0.5105	46.27	66.14
DANN (Ganin et al., 2016)		35.29	52.17	0.4978	51.19	53.19	35.76	52.68	0.4935	43.05	67.84
CLAN (Luo et al., 2019)		26.41	41.78	0.3941	35.02	51.78	34.10	50.86	0.4747	41.14	66.60
STADA (Liang et al., 2023)		5.34	10.13	0.0928	5.61	52.16	11.28	20.27	0.1854	11.59	80.64
AdaptSegNet (Tsai et al., 2018)		25.93	41.18	0.3847	37.74	45.30	30.15	46.33	0.4242	38.69	57.72
MemoryAdaptSegNet (Zhu et al., 2023)		28.89	44.82	0.4197	45.48	44.19	34.24	51.02	0.4714	45.09	58.74
Our AWDA		<b>51.85</b>	<b>68.29</b>	<b>0.6655</b>	<b>73.41</b>	<b>63.84</b>	<b>50.43</b>	<b>67.05</b>	<b>0.6478</b>	<b>54.19</b>	<b>87.93</b>

**Table 6**  
Analysis of the proposed AWDA and other comparison methods in Tables 3, 4, and 5.

Method	Paradigm	Performance	Reason
OnlySrc-HRNet, OnlySrc-SegFormer, OnlySrc	Source domain training based on different models	Low recall, high precision → low IoU/F1/Kappa	Change information is <b>weakened</b> from source to target domains due to class imbalance.
Fixed-weight OnlySrc & Decaying-weight OnlySrc	Class-weighted source domain training	More improved recall, slightly reduced precision → higher IoU/F1/Kappa	Change information is <b>enhanced</b> in target domain by source-domain class weighting training.
MMD	Source domain training + cross-domain global distribution alignment	Lower recall, slightly higher precision → lower IoU/F1/Kappa	Unstable <b>explicit</b> cross-domain alignment under the interference of class imbalance.
CLAN, STADA, AdaptSegNet, MemoryAdaptSegNet	Source domain training + iterative cross-domain adversarial training at different levels	Unstable recall and precision → unstable IoU/F1/Kappa	Unstable <b>implicit</b> cross-domain alignment under the interference of class imbalance.
DANN	Source domain training + gradient-reversal cross-domain adversarial training	Mostly increased recall and slightly reduced precision → mostly increased IoU/F1/Kappa	<b>Relatively stable</b> cross-domain alignment with smooth one-step gradient reversal operation.
FixMatch	Source domain training + target domain self-training	Mostly much lower recall, higher precision → mostly lower IoU/F1/Kappa	Class imbalance is <b>exacerbated</b> when existing domain shifts.
Our AWDA	Source domain training + cross-domain adversarial training + target domain self-training	Significantly improved recall and precision → overall increased IoU/F1/Kappa	<b>Adaptive class-balanced</b> target-domain self-training with <b>stable</b> domain shift reduction.

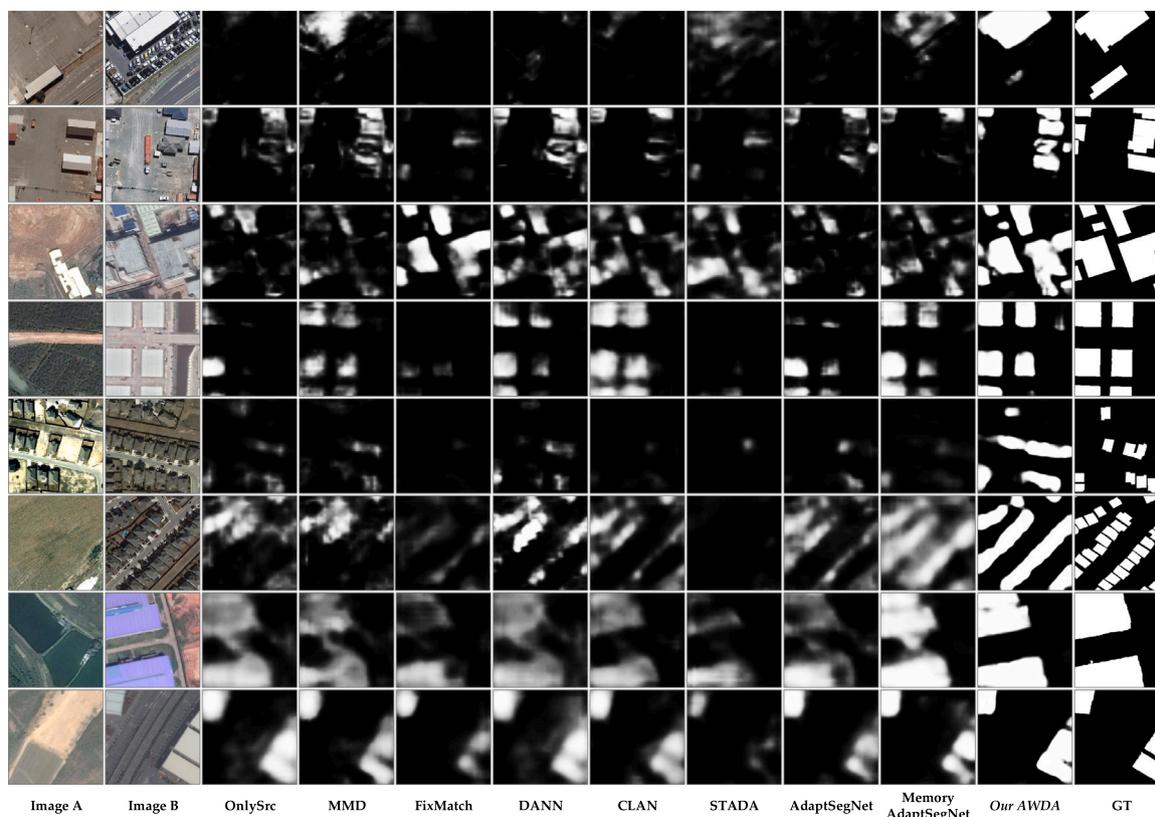


Fig. 7. Some sample visualizations obtained from different comparison methods including our AWDA. The 1–2 rows, 3–4 rows, 5–6 rows, and 7–8 rows respectively show the samples in the adaptation scenarios of LEVIR→WHU, LEVIR→GZ, WHU→LEVIR, and WHU→GZ, respectively.

predictions for change areas with low uncertainty, benefiting from its class-weighted self-training strategy. Among the comparative methods, MemoryAdaptSegNet achieves relatively recognizable predictions, however, it nonetheless exhibits challenges, including blurred edges and undetected change areas.

#### 4.8. Discussions

To further discuss experimental results of the proposed AWDA and the comparison methods in Section 4.7 as well as the source-domain weighting strategies in Section 4.6, we summarize their characteristics as shown in Table 6.

Despite the promising CD capabilities of the proposed AWDA framework, several limitations remain, with potential solutions to be explored in future work: (1) Lack of explainability in the change detection process, which could be improved through semantic-guided strategies; and (2) Blurry edge detection, particularly for buildings, which could be addressed using object-aware consistency strategies. For instance, the segmented parts obtained from Segment Anything (Kirillov et al., 2023) can be treated as the same objects to facilitate consistency learning.

## 5. Conclusion

In this paper, we introduce a novel and practical research topic termed Domain Adaptive Change Detection (DACD). DACD aims to transfer a CD model, trained on a labeled source domain (dataset), to an unlabeled target domain, thus eliminating the need for labels in the target domain. To address the challenges of domain shifts and class imbalance, which hinder the effective transfer of change knowledge, we propose a novel Adversarial and Weighted Domain Adaptation (AWDA) framework for DACD. Extensive experimental results demonstrate that the proposed AWDA framework effectively and accurately detects change areas in the target domain, significantly outperforming several advanced DA methods.

## CRedit authorship contribution statement

**Xueting Zhang:** Writing – review & editing, Writing – original draft, Methodology, Conceptualization. **Xin Huang:** Writing – review & editing, Methodology, Conceptualization. **Jiayi Li:** Writing – review & editing, Methodology, Conceptualization.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgments

The research was supported by the National Natural Science Foundation of China (under Grants 42271328).

## References

- Alesheikh, A.A., Ghorbanali, A., Nouri, N., 2007. Coastline change detection using remote sensing. *Int. J. Environ. Sci. Technol.* 4 (1), 61–66.
- Anniballe, R., Noto, F., Scalia, T., Bignami, C., Stramondo, S., Chini, M., Pierdicca, N., 2018. Earthquake damage mapping: An overall assessment of ground surveys and VHR image change detection after L'Aquila 2009 earthquake. *Remote Sens. Environ.* 210, 166–178.
- Bai, T., Wang, L., Yin, D., Sun, K., Chen, Y., Li, W., Li, D., 2023. Deep learning for change detection in remote sensing: a review. *Geo- Spat. Inf. Sci.* 26 (3), 262–288.
- Bandara, W.G.C., Patel, V.M., 2022. Revisiting consistency regularization for semi-supervised change detection in remote sensing images. *arXiv preprint arXiv:2204.08454*.
- Chen, L., Chen, H., Wei, Z., Jin, X., Tan, X., Jin, Y., Chen, E., 2022. Reusing the task-specific classifier as a discriminator: Discriminator-free adversarial domain adaptation. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 7181–7190.

- Chen, H., Qi, Z., Shi, Z., 2021. Remote sensing image change detection with transformers. *IEEE Trans. Geosci. Remote Sens.* 60, 1–14.
- Chen, H., Shi, Z., 2020. A spatial-temporal attention-based method and a new dataset for remote sensing image change detection. *Remote. Sens.* 12 (10), 1662.
- Chen, C., Xie, W., Huang, W., Rong, Y., Ding, X., Huang, Y., Xu, T., Huang, J., 2019. Progressive feature alignment for unsupervised domain adaptation. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 627–636.
- Cheng, Y., Wei, F., Bao, J., Chen, D., Zhang, W., 2023. Adpl: Adaptive dual path learning for domain adaptation of semantic segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.*
- Cubuk, E.D., Zoph, B., Shlens, J., Le, Q.V., 2020. Randaugment: Practical automated data augmentation with a reduced search space. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*. pp. 702–703.
- Daudt, R.C., Le Saux, B., Boulch, A., 2018. Fully convolutional siamese networks for change detection. In: *2018 25th IEEE International Conference on Image Processing. ICIP, IEEE*, pp. 4063–4067.
- Decuyper, M., Chávez, R.O., Lohbeck, M., Lastra, J.A., Tsendbazar, N., Hackländer, J., Herold, M., Vågen, T.-G., 2022. Continuous monitoring of forest change dynamics with satellite time series. *Remote Sens. Environ.* 269, 112829.
- Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., Fei-Fei, L., 2009. Imagenet: A large-scale hierarchical image database. In: *2009 IEEE Conference on Computer Vision and Pattern Recognition*. Ieee, pp. 248–255.
- Dong, S., Wang, L., Du, B., Meng, X., 2024. ChangeCLIP: Remote sensing change detection with multimodal vision-language representation learning. *ISPRS J. Photogramm. Remote Sens.* 208, 53–69.
- Ganin, Y., Ustinova, E., Ajakan, H., Germain, P., Larochelle, H., Laviolette, F., March, M., Lempitsky, V., 2016. Domain-adversarial training of neural networks. *J. Mach. Learn. Res.* 17 (59), 1–35.
- Giustarini, L., Hostache, R., Matgen, P., Schumann, G.J.-P., Bates, P.D., Mason, D.C., 2012. A change detection approach to flood mapping in urban areas using TerraSAR-X. *IEEE Trans. Geosci. Remote Sens.* 51 (4), 2417–2430.
- Gretton, A., Borgwardt, K., Rasch, M., Schölkopf, B., Smola, A., 2006. A kernel method for the two-sample-problem. *Adv. Neural Inf. Process. Syst.* 19.
- Gretton, A., Borgwardt, K.M., Rasch, M.J., Schölkopf, B., Smola, A., 2012a. A kernel two-sample test. *J. Mach. Learn. Res.* 13 (1), 723–773.
- Gretton, A., Sejdinovic, D., Strathmann, H., Balakrishnan, S., Pontil, M., Fukumizu, K., Sriperumbudur, B.K., 2012b. Optimal kernel choice for large-scale two-sample tests. *Adv. Neural Inf. Process. Syst.* 25.
- Guo, X., Yang, C., Li, B., Yuan, Y., 2021. Metacorection: Domain-aware meta loss correction for unsupervised domain adaptation in semantic segmentation. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 3927–3936.
- HassanPour Zonoozi, M., Seydi, V., 2023. A survey on adversarial domain adaptation. *Neural Process. Lett.* 55 (3), 2429–2469.
- He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. pp. 770–778.
- Hoyer, L., Dai, D., Van Gool, L., 2022. Daformer: Improving network architectures and training strategies for domain-adaptive semantic segmentation. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 9924–9935.
- Huang, W., Shi, Y., Xiong, Z., Wang, Q., Zhu, X.X., 2023a. Semi-supervised bidirectional alignment for remote sensing cross-domain scene classification. *ISPRS J. Photogramm. Remote Sens.* 195, 192–203.
- Huang, W., Shi, Y., Xiong, Z., Zhu, X.X., 2023b. AdaptMatch: Adaptive matching for semisupervised binary segmentation of remote sensing images. *IEEE Trans. Geosci. Remote Sens.* 61, 1–16.
- Huang, Y., Zhang, P., 2024. CSDACD: Domain-adaptive change detection network for cross-seasonal remote sensing images. *IEEE Geosci. Remote Sens. Lett.*
- Ji, S., Wei, S., Lu, M., 2018. Fully convolutional networks for multisource building extraction from an open aerial and satellite imagery data set. *IEEE Trans. Geosci. Remote Sens.* 57 (1), 574–586.
- Kang, G., Jiang, L., Yang, Y., Hauptmann, A.G., 2019. Contrastive adaptation network for unsupervised domain adaptation. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 4893–4902.
- Kennedy, R.E., Townsend, P.A., Gross, J.E., Cohen, W.B., Bolstad, P., Wang, Y., Adams, P., 2009. Remote sensing change detection tools for natural resource managers: Understanding concepts and tradeoffs in the design of landscape monitoring projects. *Remote Sens. Environ.* 113 (7), 1382–1396.
- Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A.C., Lo, W.-Y., et al., 2023. Segment anything. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*. pp. 4015–4026.
- Kou, R., Fang, B., Chen, G., Wang, L., 2020. Progressive domain adaptation for change detection using season-varying remote sensing images. *Remote. Sens.* 12 (22), 3815.
- Kumar, A., Ma, T., Liang, P., 2020. Understanding self-training for gradual domain adaptation. In: *International Conference on Machine Learning*. PMLR, pp. 5468–5479.
- Lee, J., Lee, G., 2023. Feature alignment by uncertainty and self-training for source-free unsupervised domain adaptation. *Neural Netw.* 161, 682–692.
- Li, G., Kang, G., Zhu, Y., Wei, Y., Yang, Y., 2021. Domain consensus clustering for universal domain adaptation. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 9757–9766.
- Li, L., Yang, J., Ma, Y., Kong, X., 2023. Pseudo-labeling integrating centers and samples with consistent selection mechanism for unsupervised domain adaptation. *Inform. Sci.* 628, 50–69.
- Li, Y., Yuan, L., Vasconcelos, N., 2019. Bidirectional learning for domain adaptation of semantic segmentation. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 6936–6945.
- Liang, C., Cheng, B., Xiao, B., Dong, Y., 2023. Unsupervised domain adaptation for remote sensing image segmentation based on adversarial learning and self-training. *IEEE Geosci. Remote. Sens. Lett.* 20, 1–5.
- Liang, J., Hu, D., Feng, J., 2021. Domain adaptation with auxiliary target domain-oriented classifier. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 16632–16642.
- Liu, R., Kuffer, M., Persello, C., 2019. The temporal dynamics of slums employing a CNN-based change detection approach. *Remote. Sens.* 11 (23), 2844.
- Liu, J., Xuan, W., Gan, Y., Zhan, Y., Liu, J., Du, B., 2022. An end-to-end supervised domain adaptation framework for cross-domain change detection. *Pattern Recognit.* 132, 108960.
- Long, M., Cao, Y., Wang, J., Jordan, M., 2015. Learning transferable features with deep adaptation networks. In: *International Conference on Machine Learning*. PMLR, pp. 97–105.
- Luo, Y., Zheng, L., Guan, T., Yu, J., Yang, Y., 2019. 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*. pp. 2507–2516.
- Mou, L., Bruzzone, L., Zhu, X.X., 2018. Learning spectral-spatial-temporal features via a recurrent convolutional neural network for change detection in multispectral imagery. *IEEE Trans. Geosci. Remote Sens.* 57 (2), 924–935.
- Ngondo, J., Mango, J., Liu, R., Nobert, J., Dubi, A., Cheng, H., 2021. Land-use and land-cover (LULC) change detection and the implications for coastal water resource management in the Wami-Ruvu Basin, Tanzania. *Sustainability* 13 (8), 4092.
- Ning, X., Zhang, H., Zhang, R., Huang, X., 2024. Multi-stage progressive change detection on high resolution remote sensing imagery. *ISPRS J. Photogramm. Remote Sens.* 207, 231–244.
- Pan, S.J., Tsang, I.W., Kwok, J.T., Yang, Q., 2010. Domain adaptation via transfer component analysis. *IEEE Trans. Neural Netw.* 22 (2), 199–210.
- Peng, D., Bruzzone, L., Zhang, Y., Guan, H., Ding, H., Huang, X., 2020. SemiCDNet: A semisupervised convolutional neural network for change detection in high resolution remote-sensing images. *IEEE Trans. Geosci. Remote Sens.* 59 (7), 5891–5906.
- Qing, Y., Ming, D., Wen, Q., Weng, Q., Xu, L., Chen, Y., Zhang, Y., Zeng, B., 2022. Operational earthquake-induced building damage assessment using CNN-based direct remote sensing change detection on superpixel level. *Int. J. Appl. Earth Obs. Geoinf.* 112, 102899.
- Qu, J., Dong, W., Yang, Y., Zhang, T., Li, Y., Du, Q., 2024a. Cycle-refined multidecision joint alignment network for unsupervised domain adaptive hyperspectral change detection. *IEEE Trans. Neural Netw. Learn. Syst.*
- Qu, Y., Li, J., Huang, X., Wen, D., 2023. TD-SSCD: A novel network by fusing temporal and differential information for self-supervised remote sensing image change detection. *IEEE Trans. Geosci. Remote Sens.*
- Qu, X., Liu, L., Zhu, L., Nie, L., Zhang, H., 2024b. Source-free style-diversity adversarial domain adaptation with privacy-preservation for person re-identification. *Knowl.-Based Syst.* 283, 111150.
- Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation. In: *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015: 18th International Conference, Munich, Germany, October 5–9, 2015, Proceedings, Part III* 18. Springer, pp. 234–241.
- Saito, K., Watanabe, K., Ushiku, Y., Harada, T., 2018. Maximum classifier discrepancy for unsupervised domain adaptation. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. pp. 3723–3732.
- Shafique, A., Cao, G., Khan, Z., Asad, M., Aslam, M., 2022. Deep learning-based change detection in remote sensing images: A review. *Remote. Sens.* 14 (4), 871.
- Shi, W., Caballero, J., Huszar, F., Totz, J., Aitken, A.P., Bishop, R., Rueckert, D., Wang, Z., 2016. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. pp. 1874–1883.
- Sohn, K., Berthelot, D., Carlini, N., Zhang, Z., Zhang, H., Raffel, C.A., Cubuk, E.D., Kurakin, A., Li, C.-L., 2020. Fixmatch: Simplifying semi-supervised learning with consistency and confidence. *Adv. Neural Inf. Process. Syst.* 33, 596–608.
- Tang, H., Chen, K., Jia, K., 2020. Unsupervised domain adaptation via structurally regularized deep clustering. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 8725–8735.
- Tsai, Y.-H., Hung, W.-C., Schuster, S., Sohn, K., Yang, M.-H., Chandraker, M., 2018. Learning to adapt structured output space for semantic segmentation. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. pp. 7472–7481.

- Wang, J., Sun, K., Cheng, T., Jiang, B., Deng, C., Zhao, Y., Liu, D., Mu, Y., Tan, M., Wang, X., et al., 2020. Deep high-resolution representation learning for visual recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* 43 (10), 3349–3364.
- Wang, Q., Yuan, Z., Du, Q., Li, X., 2018. GETNET: A general end-to-end 2-D CNN framework for hyperspectral image change detection. *IEEE Trans. Geosci. Remote Sens.* 57 (1), 3–13.
- Westfechtel, T., Yeh, H.-W., Zhang, D., Harada, T., 2024. Gradual source domain expansion for unsupervised domain adaptation. In: *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*. pp. 1946–1955.
- Xia, R., Zhao, C., Zheng, M., Wu, Z., Sun, Q., Tang, Y., 2023. CMDA: Cross-modality domain adaptation for nighttime semantic segmentation. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*. pp. 21572–21581.
- Xie, E., Wang, W., Yu, Z., Anandkumar, A., Alvarez, J.M., Luo, P., 2021. SegFormer: Simple and efficient design for semantic segmentation with transformers. *Adv. Neural Inf. Process. Syst.* 34, 12077–12090.
- Yuan, S., Zhong, R., Yang, C., Li, Q., Dong, Y., 2024. Dynamically updated semi-supervised change detection network combining cross-supervision and screening algorithms. *IEEE Trans. Geosci. Remote Sens.*
- Zeng, W., Cheng, M., Yuan, Z., Dai, W., Wu, Y., Liu, W., Wang, C., 2024. Domain adaptive remote sensing image semantic segmentation with prototype guidance. *Neurocomputing* 580, 127484.
- Zhan, Y., Fu, K., Yan, M., Sun, X., Wang, H., Qiu, X., 2017. Change detection based on deep siamese convolutional network for optical aerial images. *IEEE Geosci. Remote Sens. Lett.* 14 (10), 1845–1849.
- Zhang, C., Feng, Y., Hu, L., Tapete, D., Pan, L., Liang, Z., Cigna, F., Yue, P., 2022. A domain adaptation neural network for change detection with heterogeneous optical and SAR remote sensing images. *Int. J. Appl. Earth Obs. Geoinf.* 109, 102769.
- Zhang, X., Huang, X., Li, J., 2023a. Joint self-training and rebalanced consistency learning for semi-supervised change detection. *IEEE Trans. Geosci. Remote Sens.*
- Zhang, X., Huang, X., Li, J., 2023b. Semisupervised change detection with feature-prediction alignment. *IEEE Trans. Geosci. Remote Sens.* 61, 1–16.
- Zhang, Y., Wang, Z., He, W., 2023c. Class relationship embedded learning for source-free unsupervised domain adaptation. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 7619–7629.
- Zhang, C., Yue, P., Tapete, D., Jiang, L., Shanguan, B., Huang, L., Liu, G., 2020. A deeply supervised image fusion network for change detection in high resolution bi-temporal remote sensing images. *ISPRS J. Photogramm. Remote Sens.* 166, 183–200.
- Zhao, S., Li, B., Yue, X., Gu, Y., Xu, P., Hu, R., Chai, H., Keutzer, K., 2019. Multi-source domain adaptation for semantic segmentation. *Adv. Neural Inf. Process. Syst.* 32.
- Zhao, X., Mithun, N.C., Rajvanshi, A., Chiu, H.-P., Samarasekera, S., 2024. Unsupervised domain adaptation for semantic segmentation with Pseudo Label Self-Refinement. In: *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*. pp. 2399–2409.
- Zhao, H., Shi, J., Qi, X., Wang, X., Jia, J., 2017. Pyramid scene parsing network. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. pp. 2881–2890.
- Zhao, S., Yue, X., Zhang, S., Li, B., Zhao, H., Wu, B., Krishna, R., Gonzalez, J.E., Sangiovanni-Vincentelli, A.L., Seshia, S.A., et al., 2020. A review of single-source deep unsupervised visual domain adaptation. *IEEE Trans. Neural Netw. Learn. Syst.* 33 (2), 473–493.
- Zheng, Z., Wan, Y., Zhang, Y., Xiang, S., Peng, D., Zhang, B., 2021. CLNet: Cross-layer convolutional neural network for change detection in optical remote sensing imagery. *ISPRS J. Photogramm. Remote Sens.* 175, 247–267.
- Zhou, L., Li, N., Ye, M., Zhu, X., Tang, S., 2024. Source-free domain adaptation with class prototype discovery. *Pattern Recognit.* 145, 109974.
- Zhu, J., Guo, Y., Sun, G., Yang, L., Deng, M., Chen, J., 2023. Unsupervised domain adaptation semantic segmentation of high-resolution remote sensing imagery with invariant domain-level prototype memory. *IEEE Trans. Geosci. Remote Sens.* 61, 1–18.
- Zou, Y., Yu, Z., Kumar, B., Wang, J., 2018. Unsupervised domain adaptation for semantic segmentation via class-balanced self-training. In: *Proceedings of the European Conference on Computer Vision. ECCV*, pp. 289–305.
- Zuo, Y., Li, L., Liu, X., Gao, Z., Jiao, L., Liu, F., Yang, S., 2024. Robust instance-based semi-supervised learning change detection for remote sensing images. *IEEE Trans. Geosci. Remote Sens.*