Article

Global and Local Graph-Based Difference Image Enhancement for Change Detection

Xiaolong Zheng, Dongdong Guan *, Bangjie Li, Zhengsheng Chen and Lefei Pan

High-Tech Institute of Xi'an, Xi'an 710025, China
* Correspondence: guandd@hotmail.com

Abstract: Change detection (CD) is an important research topic in remote sensing, which has been applied in many fields. In the paper, we focus on the post-processing of difference images (DIs), i.e., how to further improve the quality of a DI after the initial DI is obtained. The importance of DIs for CD problems cannot be overstated, however few methods have been investigated so far for re-processing DIs after their acquisition. In order to improve the DI quality, we propose a global and local graph-based DI-enhancement method (GLGDE) specifically for CD problems; this is a plug-and-play method that can be applied to both homogeneous and heterogeneous CD. GLGDE first segments the multi-temporal images and DIs into superpixels with the same boundaries and then constructs two graphs for the DI with superpixels as vertices: one is the global feature graph that characterizes the association between the similarity relationships of connected vertices in the multi-temporal images and their changing states in a DI, the other is the local spatial graph that exploits the change information and contextual information of the DI. Based on these two graphs, a DI-enhancement model is built, which constrains the enhanced DI to be smooth on both graphs. Therefore, the proposed GLGDE can not only smooth the DI but also correct the it. By solving the minimization model, we can obtain an improved DI. The experimental results and comparisons on different CD tasks with six real datasets demonstrate the effectiveness of the proposed method.

Keywords: change detection; difference image; smoothness; graph; heterogeneous data

1. Introduction

1.1. Background

The change detection (CD) of remote sensing images is a technique to extract change information by comparing multi-temporal images acquired over the same geographical area but at different times [1]. Earth observation technology can provide long-term, wide-area, periodic observations of the Earth's surface; CD is also one of the earliest and most widely used research areas of remote sensing [2], which has also achieved very successful applications in many practical tasks, such as environmental monitoring [3], agricultural surveys [4], urban studies [5], and disaster assessment [6].

Up to now, numerous CD algorithms have been investigated and one can refer to the latest review articles [7–10]. Generally, there are different ways to classify CD algorithms. For example, (1) according to the detection and analysis granularity, it can be divided into pixel-level-, object-level-, and scene-level-based CD; (2) according to the type of output result, it can be divided into binary and multiple CD; (3) depending on the usage of label data, it can be classified as supervised, semi-supervised, and unsupervised CD; (4) according to whether deep neural networks are used, it can be divided into traditional methods and deep learning-based methods; (5) depending on the source of the input images, it can be divided into homogeneous and heterogeneous CD; (6) and, according to the technology used, it can be also divided into spectral change identification methods, post-classification comparison methods, direct multi-temporal classification methods, and

hybrid methods. In this paper, we focus on the homogeneous CD of optical images, homogeneous CD of synthetic aperture radar (SAR) images, and heterogeneous CD.

The CD process can usually be divided into three main sequential steps [11,12]: image preprocessing, differential image (DI) generation, and change extraction by analyzing the DI. In the preprocessing, radiometric correction and geometric coregistration are usually required, which enables pairwise images to correspond to the same geographic area. Then, it compares the two aligned images to obtain the DI in the second process, which is aimed at enhancing the contrast between the changed and unchanged regions. Finally, it classifies the DI to extract the changed region and obtain the change map (CM) in the third process.

DI generation has a significant impact in the CD process for the following reasons. (1) DIs have a direct influence on the accuracy of CM. A high-quality DI is highly discriminative between the changed and unchanged classes and, thus, only a simple threshold segmentation or clustering is needed to classify the DIs for obtaining CM [13–15]. (2) DIs provide richer change information. Compared to the binary CM, which can only provide certain change/unchange information, DIs are able to provide the probability of change. For some difficult problems or problems that must be handled with caution, such DIs that provide uncertainty results may be more valuable for assisting experts [11]. (3) DIs could be used to support other unsupervised CD methods. For example, DIs could assist the training process or build pseudo-training labels for some deep learning-based methods [16,17]. Therefore, we focus on how to improve the quality of DIs in the paper, while, for the final CM, we only use a simple thresholding method on DIs to obtain it.

1.2. Related Work

Here, we review some methods for obtaining DIs in different CD tasks. Define **X** and **Y** to be the multi-temporal images to be compared and let x and y be the data extracted from the same positions of **X** and **Y**, respectively, which can be individual pixels, square patches or superpixels, depending on the granularity of the detection analysis. According to the sources of **X** and **Y**, we review some relevant algorithms for calculating DIs.

1.2.1. DI of Homogeneous Optical Images

For the homogeneous CD of optical images, the image differencing (e.g., $\|x - y\|$) is the simplest way to calculate DIs, which is based on the assumption that the noise in optical images is usually additive. Change vector analysis (CVA) [18] calculates the change vectors by comparing the spectra of two images, which can provide both change magnitude and change direction. Based on the canonical correlations analysis, the multivariate alteration detection method (MAD) [19] and iteratively reweighted MAD method (IRMAD) [20] have been proposed, which is invariant to separate linear (affine) transformations of the spectra. Therefore, they can significantly reduce false alarms due to differences in device gain or linear radiation and atmospheric correction schemes. Lv et al. [15] have proposed an adaptive spatial–contextual extraction method (ASE) for CD of very high resolution optical images, which first adaptively selects a suitable local area for each pixel to exploit the contextual information and then uses a defined band-to-band (B2B) distance metric to calculate the change magnitude. By exploiting the convolutional neural network (CNN), Saha et al. [21] have combined CVA with deep neural networks and proposed deep change vector analysis (DCVA), which calculates the deep change vectors and extracts richer contextual information by utilizing the features from different layers of CNN. Finally, the binary and multiple CM are computed by analyzing the deep change vector. Du et al. have extended the slow feature analysis (SFA) [22] to the deep network and proposed a deep slow feature analysis (DSFA) for unsupervised CD [23]. DSFA first uses the CVA to find unchanged pixels to construct a pseudo-training set and then utilizes two symmetric deep networks to project the multi-temporal images into the latent feature space, where SFA is employed to identify changed and unchanged regions, and, finally, the DIs are computed using the chi-square distance metric and the CM is obtained using the thresholding method.

1.2.2. DIs of Homogeneous SAR Images

For the homogeneous CD of SAR images, it is difficult to detect the changes using an image differencing operator due to the inherent multiplicative speckle noise of SAR images. Alternatively, the ratio [24], log-ratio [25], and mean-ratio [26] operators are the commonly used DI calculation methods. Nar et al. [27] have proposed a sparsity-driven change detection (SDCD) method, which can reduce the influence of speckle noise on the DI by using an ℓ_1-norm-based total variation regularization. Sun et al. [28] have proposed a nonlocal low-rank- and a two-level clustering-based method, which jointly employs the nonlocal despeckling method for computing the DI and the cascade clustering strategy for calculating the binary CM. In [29], an adaptive Contourlet fusion clustering (CFC)-based CD method has been proposed, which combines the log-ratio and mean-ratio images to generate a new DI by using the adaptive Contourlet fusion and then segments the fused DI by using a fast non-local clustering algorithm to reduce the impact of the noise. A hierarchical heterogeneous graph-driven CD method has been proposed in [30], which combines two inter-connected pixel-based and superpixel-based graph layers to fully exploit the structure information of images and then uses the graph cuts to separate the DI to obtain CM. A log cumulants and stacked autoencoder-based method has been proposed for detecting changes caused by fires using SAR images acquired by Sentinel-1 in [31], which extracts the features by using a tunable Q wavelet transformation with higher order log-cumulants statistics and uses the stacked autoencoder to classify the changed and unchanged regions. Zhang et al. [32] have proposed a spatial–temporal gray-level co-occurrence aware network (STGCNet) to suppress the influence of speckles on detecting changes, which first uses the log-ratio operation to generate an initial DI and uses the fuzzy c-means (FCM) clustering to select the reliable changed and unchanged training samples and then trains the two-stream STGCNet to mine the spatial–temporal information. A dynamic graph-level neural network (DGLNN) is proposed in [33], which builds a dynamic graph on the three-channel pixel neighborhood block constructed from the initial log-ratio-based DI and multi-temporal SAR images and learns discriminative representations for each block by using feature propagation and node aggregation in the graph.

1.2.3. DI of Heterogeneous CD

For the heterogeneous CD, the compared multi-temporal images are acquired by different sensors and characterize different physical quantities, so it is impossible to obtain the DI by directly comparing the heterogeneous images. Therefore, it is essential to transform heterogenous images into the same metric space for comparison. Liu et al. [13] have proposed the homogeneous pixel transformation (HPT) method, which uses the labeled unchanged pixel pairs to transform one image (e.g., $\mathbf{X}$) to the domain of the other images ($\mathbf{Y}$) with kernel regression functions and then compare the regression image ($\mathbf{Y}'$) and original image ($\mathbf{Y}$) in the same domain using image differencing ($\|\mathbf{Y} - \mathbf{Y}'\|$). To overcome the dependence on labeled samples, some unsupervised regression methods have also been proposed for heterogeneous CD. In [14], an affinity matrix distance (AMD) is used to pick samples that have a high probability of being unchanged and then four traditional regression functions are trained based on the pseudo-training set. Furthermore, the AMD has also been used to train deep regression networks as the prior change not only to construct training samples [17] but also to assist in the training process [16]. Mignotte proposed a fractal projection and Markovian segmentation-based algorithm (FPMS) for heterogeneous CD [34], which consists of a fractal encoding step that encodes the pre-event image and a fractal decoding step that projects the pre-event image to the domain of a post-event image and then computes and segments the DI using Markovian segmentation algorithms. Based on the structure consistency, Sun et al. proposed several graph-based heterogeneous CD methods. For example, they proposed the structure regression-based methods that constrain the structures of original images and regression images that are the same and causes the changed image to be sparse [35–37]; they also proposed the structure comparison-based methods that first construct two graphs of heterogeneous images to

capture the structure information and then compare the structures using graph projection [11,38,39]. Recently, they have also proposed a heterogeneous CD framework based on graph signal processing [40] and analyzed the heterogeneous CD problem from the vertex and spectral domains of the graph, respectively.

According to the introduction of the above-related methods, we have the following findings.

- Different CD problems face different challenges. (1) For the CD of homogeneous optical images, its difficulty lies in that, when the image resolution is very high, the great intraclass variation and low interclass variance as well as the influence of illuminations and seasons can lead to a lot of salt-and-pepper noise [7]. (2) For the CD of homogeneous SAR images, its difficulty lies in the inherent speckle noise and high intensity variation that can lead to difficult trade-offs between noise removal and geometrical detail preservation in the DI. (3) For the heterogeneous CD, the key lies in how to construct relationships between heterogeneous images so that incomparable images can be compared; it also faces the challenges of both the homogeneous CD of optical images and SAR images.
- How to obtain high-quality DI is one of the keys to the CD problem. After generating the difference maps, most of these methods treat them as conventional image segmentation problems to obtain the final CM, such as [13–15,17,21,27,39].

1.3. Motivations

Notwithstanding that the DI quality is very important for the performance of CD, few methods have been investigated so far for re-processing DI after its acquisition. The benefits of post-processing the DI are self-evident, as it can further enhance the quality of the DI, either for the next step of computing CM with segmentation or for assisting other methods.

In order to reduce the influence of speckle noise on DIs of homogeneous SAR images, Zheng et al. [41] have combined the mean filter and the median filter to obtain a better DI, which uses the mean filter on the DI computed using the image differencing operator for smoothing and uses the median filter on the DI computed using the log-ratio operator for preserving edge information. In [42], the DIs generated using a Gauss-log ratio, the log-ratio operators are fused using a discrete wavelet transform, and then the fused DI is filtered using the nonsubsampled contourlet transform model, which can reduce the noise of the DI and keep the edge information of the changed regions. In [14,17], after obtaining the DI, the authors have removed the outliers in DI, i.e., clipping the pixels whose values are beyond a few standard deviations from the mean value of DI, and then used the fully connected conditional random field model (CRF) proposed in [43] to filter the DI, which exploits the spatial context information to improve the quality of the DI. By drawing on the convolution and pooling operations in CNN, Zhang et al. [44] have proposed a weighted average filter for the DI generated using a log-ratio operator, which can suppress the speckle influence and enhance the edges of the DI. A graph signal smoothness representation method has been proposed in [45], which uses the smoothing property of the changed signal on the fused graph to smooth the DI.

Although the above methods can improve the quality of the DI to some extent, especially in reducing the noise influence, they still have the following two shortcomings.

- Most of these methods are for the conventional denoising and smoothing of DI and they only exploit the information of the DI itself, such as the change information (pixel value) and spatial context information, while ignoring the specificity of the change detection task and neglecting the information in the original multi-temporal images, which limits their performance.
- Most of the methods only serve as "icing on the cake" for smoothing the DI, but cannot further correct the DI. For example, when there is an overall error in the local area in the DI, i.e., when the pixel values of the entire local area that really changed are all 0 in

the DI or when the pixel values of the entire local area that really unchanged are all 1, it is difficult to correct this error based on the spatial smoothing or filtering operations.

In order to address the above challenges, in this paper, we propose a global and local graph-based DI-enhancement method (GLGDE) for CD problems. It is a plug-and-play approach for the post-processing of DI, which can be applied to the homogeneous CD of optical images and SAR images and heterogeneous CD. Specifically, once the initial DI is obtained, e.g., a coarse DI is obtained by using the methods introduced in the related work, we smooth and correct the DI by the following steps. Firstly, we co-segment the DI and multi-temporal images into superpixels with the same shape and boundaries and then extract the features of the images. Second, we construct two graphs for the DI with superpixels as vertices; one connects each vertex and its similar vertices in the multi-temporal images to exploit the association between the similarity relationships of the multi-temporal images and the changing states of the DI and the other connects each vertex and its neighboring vertices in the local spatial neighborhood to capture the local structure information of the DI. In the constructed graphs, the edge weights not only exploit the change information and contextual information of the DI but also the correlation information of the original multi-temporal images, which is crucial for correcting and smoothing DIs. Third, based on the constructed graphs, we propose a DI-enhancement model that contains three constraint terms: a global feature graph-smoothing regularization term, a local spatial graph-smoothing regularization term, and a change data-regularization term. By solving the minimization model, we can obtain the improved DI. Finally, the binary CM can be easily obtained using threshold segmentation.

1.4. Contributions

The main contributions of this paper can be summarized as follows.

- First, we have designed a DI-enhancement algorithm specifically for the change detection task, which is a plug-and-play approach for DI post-processing. This is a rarely found work specifically designed for smoothing and correcting DIs in CD problems.
- Second, the proposed DI-enhancement algorithm, named GLGDE for short, not only can smooth the DI but also correct it by using the constructed global feature graph and local spatial graph, which can fully fuse and utilize the change and contextual information in the DI and correlation information in the multi-temporal images.
- Third, due to using superpixels as vertices, the scale of the model is small. The algorithm achieves DI improvement with low computational complexity, which would be of great practical value. Extensive experiments in different CD scenarios, i.e., homogeneous CD of SAR and optical images and heterogeneous CD, demonstrate the effectiveness of the proposed method.

1.5. Outline

The rest of this paper is structured as follows: Section 2 describes the proposed global and local graph-based DI smoothing and correction method, and Section 3 provides the experimental results. Finally, Section 4 concludes this paper.

2. Global and Local Graph-Based DI Enhancement

With two co-registered multi-temporal images acquired using the same (homogeneous) or different (heterogeneous) sensors, denoted as $\mathbf{X} \in \mathbb{R}^{M \times N \times C_x}$ and $\mathbf{Y} \in \mathbb{R}^{M \times N \times C_y}$, and the initial coarse DI obtained from other methods, denoted as $\mathbf{D} \in \mathbb{R}^{M \times N}$, the goal of this paper is to improve the quality of the DI to obtain an enhanced DI, which has a great discrimination between changed and unchanged classes. We denote the pixels of the images as $x(m, n, c)$, $y(m, n, c)$ and $d(m, n)$, respectively, and suppose that the DI is normalized, i.e., $0 \leq d(m, n) \leq 1$, and that a larger value of $d(m, n)$ indicates a higher probability of change in the region represented by (m, n).

Next, we describes the proposed DI-enhancement method in detail, which contains three main steps: pre-processing, constructing the global and local graphs, and solving the GLGDE model. Figure 1 shows the framework of the proposed DI-enhancement method.
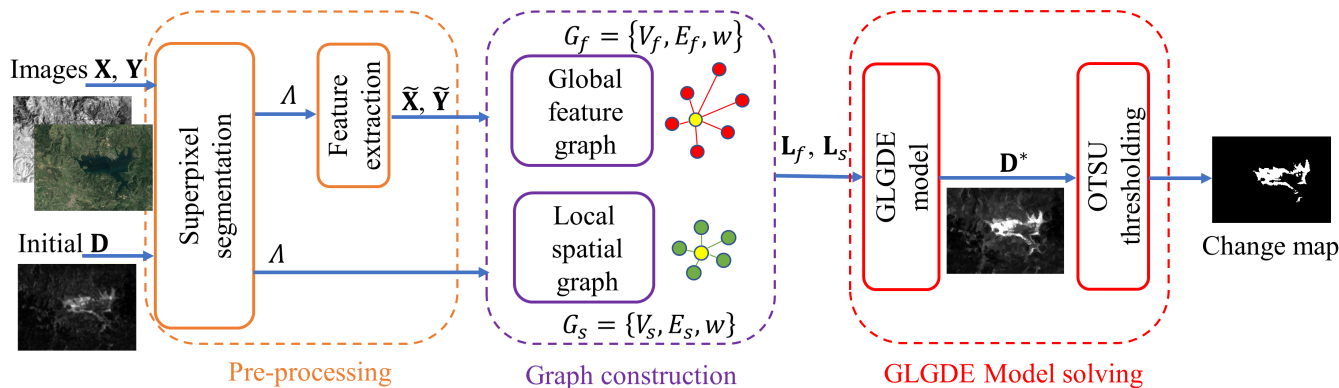

Figure 1. Framework of the proposed GLGDE.

2.1. Pre-Processing

As aforementioned, we use the graphs to capture the structure information of images. We choose to use superpixels instead of individual pixels or square patches as graph vertices, which provides two benefits. First, the size of the graph is greatly reduced, e.g., for a 1000×1000 image with individual pixels as vertices, the size of the graph is 10^6, which results in a heavy computational burden for the algorithm. On the contrary, when using superpixels as vertices, the size of the graph is only related to the number of superpixels, which will greatly reduce the complexity of the algorithm. Second, the superpixel contains rich contextual information and is able to preserve the shape and edges of the object, which is very important for the CD problem with very-high-resolution images.

In order to segment multiple images (i.e., $\mathbf{X}$, $\mathbf{Y}$, and $\mathbf{D}$) into superpixels with the same boundaries, we first construct a false RGB image $\mathbf{Z} \in \mathbb{R}^{M \times N \times 3}$. The first and second channels of $\mathbf{Z}$ are the normalized grayscale images of $\mathbf{X}$ and $\mathbf{Y}$, respectively, and the third channel of $\mathbf{Z}$ is the initial DI of $\mathbf{D}$. Note that if there is an SAR image in the multi-temporal images, e.g., $\mathbf{X}$, the corresponding channel of $\mathbf{Z}$ is the normalized log-transformed SAR image, which can convert the multiplicative noise of the SAR image into additive noise, thus facilitating the distance calculation in the subsequent superpixel segmentation. Then, the Gaussian mixture model-based superpixel segmentation method (GMMSP) [46] is employed to segment $\mathbf{Z}$ into N_S regions, in which each superpixel is associated with a weighted sum of Gaussian functions. GMMSP can efficiently generate superpixels that adhere to the boundaries of an object. We define the superpixel co-segmentation map as Λ, then we can obtain the segmented superpixels of $\mathbf{X}_i$, $\mathbf{Y}_i$, and $\mathbf{D}_i$ by projecting the map Λ to the multiple images (i.e., $\mathbf{X}$, $\mathbf{Y}$, and $\mathbf{D}$) as:

$$
\begin{aligned}
\mathbf{X}_i &= \{x(m,n,c) | (m,n) \in \Lambda_i, c = 1, \cdots, C_x\} \\
\mathbf{Y}_i &= \{y(m,n,c) | (m,n) \in \Lambda_i, c = 1, \cdots, C_y\} \\
\mathbf{D}_i &= \{d(m,n) | (m,n) \in \Lambda_i\}
\end{aligned}
\tag{1}
$$

Then, $\mathbf{X}_i$, $\mathbf{Y}_i$, and $\mathbf{D}_i$ represent the same region. More importantly, the pixels inside each superpixel obtained through co-segmentation are mostly internally homogeneous, i.e., they represent the same kind of objects, so the pixels inside each superpixel are likely to belong to the same class of labels, i.e., all changed or unchanged. In this way, the computational complexity can be reduced by using superpixels and the interference caused by inconsistent internal pixel labels can be avoided by using contextual information.

Once the co-segmentation superpixels are obtained, different features can be extracted to capture the different information of superpixel, such as the intensity (spectra) and textual

and structure information. In this paper, we simply use the mean, median and variance from each band of $\mathbf{X}_i$ and $\mathbf{Y}_i$ as the superpixel feature. Naturally, some other features are also available. Then, we can obtain the feature matrices of the images $\mathbf{X}$ and $\mathbf{Y}$, denoted as $\tilde{\mathbf{X}}$ and $\tilde{\mathbf{Y}}$, respectively, where each column corresponds to the superpixel feature vector. We define the index set as $\mathcal{I} = \{1, \cdots, N_S\}$, the feature distance between $\mathbf{X}_i$ and $\mathbf{X}_j$ as $dist_{i,j}^x = \left\| \tilde{\mathbf{X}}_i - \tilde{\mathbf{X}}_j \right\|_2^2$, and the feature distance between $\mathbf{Y}_i$ and $\mathbf{Y}_j$ as $dist_{i,j}^y = \left\| \tilde{\mathbf{Y}}_i - \tilde{\mathbf{Y}}_j \right\|_2^2$, we denote the i-th distance vectors as $dist_i^x = \left\{ dist_{i,j}^x | j \in \mathcal{I} \right\}$ and $dist_i^y = \left\{ dist_{i,j}^y | j \in \mathcal{I} \right\}$, and we define the labels of the i-th superpixel as L_i, with $L_i = 1$ indicating changed and $L_i = 0$ indicating unchanged.

2.2. Global Feature Graph

We construct two graphs for the DI, which can capture the change information and contextual information in the DI and utilize the relationships between the multi-temporal images to enhance the DI.

We define the global feature graph as $G_f = \left\{ V_f, E_f, w \right\}$, which sets each superpixel of DI as the vertex, i.e., $V_f = \{ \mathbf{D}_i | i = 1, \cdots, N_S \}$, and connect each superpixel with the superpixel corresponding to its K-nearest neighbor (KNN) in the original multi-temporal images, that is, for the i-th and j-th superpixels of $\mathbf{D}_i$ and $\mathbf{D}_j$; if $\mathbf{X}_j$ is the KNN of $\mathbf{X}_i$ or $\mathbf{Y}_j$ is the KNN of $\mathbf{Y}_i$, then $\mathbf{D}_i$ and $\mathbf{D}_j$ are connected by the edge E_f, defined as:

$$E_f = \left\{ (i,j) | i \in \mathcal{I}, j \in \mathcal{N}_i^x \cup \mathcal{N}_i^y \right\} \tag{2}$$

where $\mathcal{N}_i^x$ and $\mathcal{N}_i^y$ denote the position sets of KNN of $\mathbf{X}_i$ and $\mathbf{Y}_i$, respectively. Furthermore, the KNN sets of $\mathcal{N}_i^x$ and $\mathcal{N}_i^y$ are computed as: if and only if $dist_{i,j}^x$ belongs to one of the K-smallest elements of the distance vectors of $dist_i^x$ or $dist_j^x$, then $j \in \mathcal{N}_i^x$; for the $\mathcal{N}_i^y$, we have a similar definition. Next, we investigate the association between the similarity relationships of connected vertices in the multi-temporal images and their changing states in DI.

First, we consider the i-th vertex and j-th vertex using $j \in \mathcal{N}_i^x$. Since $\mathbf{X}_j$ is the KNN of $\mathbf{X}_i$, we have that $\mathbf{X}_i$ and $\mathbf{X}_j$ have a high probability of belonging to the same kind of object, e.g., *road*. However, for the corresponding $\mathbf{Y}_i$ and $\mathbf{Y}_j$ in the post-event image connected by this edge $(i, j), j \in \mathcal{N}_i^x$, we have two different states.

State #1, if $\mathbf{Y}_i$ and $\mathbf{Y}_j$ also belong to the same kind of object, then we have that the labels of the i-th and j-th superpixels should be same. That is, $L_i = L_j = 1$ when $\mathbf{Y}_i$ and $\mathbf{Y}_j$ also belong to the same kind of object as $\mathbf{X}_i$, e.g., *road*, and $L_i = L_j = 0$, when $\mathbf{Y}_i$ and $\mathbf{Y}_j$ belong to the same kind of object that is different from $\mathbf{X}_i$, e.g., *water*. In state #1, we have $L_i = L_j$, that is, the pixel values of $\mathbf{D}_i$ and $\mathbf{D}_j$ in the DI should be very close (all large or all small).

State #2, if $\mathbf{Y}_i$ and $\mathbf{Y}_j$ belong to the different kinds of objects, then we have that the labels of the i-th and j-th superpixels should be different. That is, (1) $L_i = 1, L_j = 0$ when $\mathbf{Y}_i$ belongs to the object of the same kind as $\mathbf{X}_i$ while $\mathbf{Y}_j$ belongs to the object of the different kind as $\mathbf{X}_i$ (e.g., *water*); (2) $L_i = 1, L_j = 0$ when $\mathbf{Y}_i$ belongs to the object of the different kinds as $\mathbf{X}_i$ (e.g., *water*) while $\mathbf{Y}_j$ belongs to the object of the same kind as $\mathbf{X}_i$. In state #2, we have $L_i \neq L_j$, that is, the pixel values of $\mathbf{D}_i$ and $\mathbf{D}_j$ in the DI should be very different (one is large, the other is small).

We define p_i as the probability that the i-th superpixel belongs to the changed class. In these two states, we can find that the value of $(p_i - p_j)^2, j \in \mathcal{N}_i^x$ is determined by the probability of whether $\mathbf{Y}_i$ and $\mathbf{Y}_j, j \in \mathcal{N}_i^x$ belong to the same kind of object. The latter can be measured using the following Gaussian kernel function as:

$$f_{i,j}^y = \exp\left(-2dist_{i,j}^y + \min_{m \in \mathcal{N}_i^y}\left\{ dist_{i,m}^y \right\} + \min_{m \in \mathcal{N}_j^y}\left\{ dist_{j,m}^y \right\} \right), j \in \mathcal{N}_i^x \tag{3}$$

Then, we have that the larger the value of $f_{i,j}^y$, the more likely it is that $\mathbf{Y}_i$ and $\mathbf{Y}_j$ belong to the same kind of object, the smaller the distance of $(p_i - p_j)^2$. On the contrary, the smaller the value of $f_{i,j}^y$, the more likely it is that $\mathbf{Y}_i$ and $\mathbf{Y}_j$ represent different kinds of objects, the larger the distance of $(p_i - p_j)^2$.

Second, we consider the i-th vertex and the j-th vertex with $j \in \mathcal{N}_i^y$. Since $\mathbf{Y}_j$ is the KNN of $\mathbf{Y}_i$, we have that $\mathbf{Y}_i$ and $\mathbf{Y}_j$ have a high probability of belonging to the same kind of object. However, for the corresponding $\mathbf{X}_i$ and $\mathbf{X}_j$ in the pre-event image connected by this this edge (i,j), $j \in \mathcal{N}_i^y$, we also have two different states, similarly to state #1 and state #2.

State #3, if $\mathbf{X}_i$ and $\mathbf{X}_j$ also belong to the same kind of object, then we have that the labels of the i-th and j-th superpixels should be same, i.e., $L_i = L_j$.

State #4, if $\mathbf{X}_i$ and $\mathbf{X}_j$ belong to the different kinds of objects, then we have that the labels of the i-th and j-th superpixels should be different, i.e., $L_i \neq L_j$.

Similarly, the value of $(p_i - p_j)^2$ is determined by the probability of whether $\mathbf{X}_i$ and $\mathbf{X}_j$ belong to the same kind of object, which can be measured using the following function:

$$f_{i,j}^x = \exp\left(-2dist_{i,j}^x + \min_{m \in \mathcal{N}_i^x}\{dist_{i,m}^x\} + \min_{m \in \mathcal{N}_j^x}\{dist_{j,m}^x\}\right), \ j \in \mathcal{N}_i^y \qquad (4)$$

We also have that the larger the value of $f_{i,j}^x$, the smaller the distance of $(p_i - p_j)^2$ and the smaller the value of $f_{i,j}^x$, the larger the distance of $(p_i - p_j)^2$.

By using Equations (3) and (4), we can set the edge weight of the graph G_f as:

$$W_{i,j}^f = f_{i,j}^x \delta\left(j \in \mathcal{N}_i^y\right) + f_{i,j}^y \delta(j \in \mathcal{N}_i^x), \forall(i,j) \in E_f \qquad (5)$$

where $\delta(\cdot)$ is the discriminant function: when the condition in parentheses holds, it adopts the value of 1; otherwise, it adopts the value of 0.

In the graph G_f, the i-th and j-th superpixels of $\mathbf{D}_i$ and $\mathbf{D}_j$ are connected by the edge of $(i,j) \in E_f$ with the weight $W_{i,j}^f$ of (5). We can find that the edges are constructed in the global feature space of the multi-temporal images and the weights are determined based on the relationships between the multi-temporal images.

Based on the above analysis, we have the following regularization:

$$\min_{\mathbf{p} \in \mathbb{R}^{N_S}} \sum_{(i,j) \in E_f} W_{i,j}^f (p_i - p_j)^2 \qquad (6)$$

By defining the Laplacian matrix of graph G_f as $\mathbf{L}_f$, the regularization (6) can be rewritten as:

$$\min_{\mathbf{p} \in \mathbb{R}^{N_S}} 2\mathbf{p}^T \mathbf{L}_f \mathbf{p} \qquad (7)$$

Then, we can find that it requires the DI to be smooth on the graph G_f. This global feature graph-induced change-smoothness-based regularization has two advantages: first, it characterizes the association between the similarity relationships of the connected vertices in the multi-temporal images and their changing states in DI, which can be used to smooth the DI globally, i.e., to correct the DI, as it exploits information in the global feature space rather than the local spatial information. Second, this change smoothness is widespread because we make no other qualifying assumptions on the CD problem in the derivation, so it can be applied to the both homogeneous and heterogeneous CD problems and it can also be applied to other methods as a constraint on the change, such as the image regression-based CD methods [35,36,47].

2.3. Local Spatial Graph

We define the local spatial graph as $G_s = \{V_s, E_s, w\}$, which sets each superpixel of the DI as the vertex, i.e., $V_s = \{\mathbf{D}_i | i = 1, \cdots, N_S\}$, and connected each superpixel with

its spatially close neighbors, defined as *R*-adjacent neighbors. That is, two superpixels of $\mathbf{D}_i$ and $\mathbf{D}_j$ are connected by the graph G_s as long as their boundaries intersect or the distance between their center points is less than *R*, denote as $i \in \mathcal{N}_j^R$ or $j \in \mathcal{N}_i^R$. Since the average pixel number of the co-segmented superpixels is MN/N_S, we set $R = 2\sqrt{MN/N_S}$ for simplicity. Then, we have:

$$E_s = \left\{ (i,j) | i \in \mathcal{I}, j \in \mathcal{N}_i^R \right\} \tag{8}$$

Inspired by the *n*-links proposed in [38], we set the weight of $W_{i,j}^s$ for the graph G_s as:

$$W_{i,j}^s = \frac{1}{dist_{i,j}^\Lambda} \times \begin{cases} \exp\left(-\frac{dist_{i,j}^y}{2\sigma_1^2} - \frac{dist_{i,j}^x}{2\sigma_2^2}\right), & \text{if } dist_{i,j}^y \leq \sigma_1^2, dist_{i,j}^x \leq \sigma_2^2 \\ \exp\left(\frac{dist_{i,j}^y}{2\sigma_1^2} - \frac{dist_{i,j}^x}{2\sigma_2^2} - 1\right), & \text{if } dist_{i,j}^y \leq \sigma_1^2, dist_{i,j}^x > \sigma_2^2 \\ \exp\left(-\frac{dist_{i,j}^y}{2\sigma_1^2} + \frac{dist_{i,j}^x}{2\sigma_2^2} - 1\right), & \text{if } dist_{i,j}^y > \sigma_1^2, dist_{i,j}^x \leq \sigma_2^2 \\ \exp(-1), & \text{if } dist_{i,j}^y > \sigma_1^2, dist_{i,j}^x > \sigma_2^2 \end{cases} \tag{9}$$

where $dist_{i,j}^\Lambda$ denote the spatial distance between the center points of Λ_i and Λ_j, $\sigma_1^2 = \left\langle dist_{i,j}^y \right\rangle$ and $\sigma_2^2 = \left\langle dist_{i,j}^x \right\rangle$ are the two normalized parameters using $\langle \cdot \rangle$ to calculate the mean value of the feature distance throughout the entire image. From (9), there are four different cases of edge weight assignment: a large weight when $\mathbf{X}_i$ and $\mathbf{X}_j$ and $\mathbf{Y}_i$ and $\mathbf{Y}_j$ are both similar; a small weight when $\mathbf{X}_i$ and $\mathbf{X}_j$ are not similar but $\mathbf{Y}_i$ and $\mathbf{Y}_j$ are similar; a small weight when $\mathbf{X}_i$ and $\mathbf{X}_j$ are similar but $\mathbf{Y}_i$ and $\mathbf{Y}_j$ are not similar; and a median weight when $\mathbf{X}_i$ and $\mathbf{X}_j$ and $\mathbf{Y}_i$ and $\mathbf{Y}_j$ are both not similar.

With this defined edge weight $W_{i,j}^s$, the constructed local spatial graph G_s not only exploits the contextual information of the DI but also the similarity information of the multi-temporal image. Based on spatial continuity, we have the following regularization:

$$\min_{\mathbf{p} \in \mathbb{R}^{N_S}} \sum_{(i,j) \in E_s} W_{i,j}^s \left(p_i - p_j \right)^2 \tag{10}$$

From (10), we can find that for the spatially adjacent $\mathbf{D}_i$ and $\mathbf{D}_j$, the regularization (10) requires that their values are close, especially for those that originally have the same similarity relationship in the two multi-temporal images. By defining the Laplacian matrix of graph G_s as $\mathbf{L}_s$, the regularization (10) can be rewritten as:

$$\min_{\mathbf{p} \in \mathbb{R}^{N_S}} 2\mathbf{p}^T \mathbf{L}_s \mathbf{p} \tag{11}$$

It can be regarded as a local spatial smoothness constraint for the DI.

2.4. GLGDE Model

For the DI-enhancement model, we have a change data regularization term, which constrains the difference between the enhanced DI and the original DI, defined as:

$$\min_{\mathbf{p} \in \mathbb{R}^{N_S}} \left\| \mathbf{p} - \bar{\mathbf{d}} \right\|_2^2 \tag{12}$$

where $\bar{\mathbf{d}}$ is the mean vector of DI with $\bar{d}_i$ being the mean value of the superpixel $\mathbf{D}_i$ as $\bar{d}_i = \frac{\sum_{(m,n) \in \Lambda_i} d(m,n)}{|\Lambda_i|}$.

By combining the global feature graph G_f-induced regularization (GRGR) of (7), the local spatial graph G_s-induced regularization (LSGR) of (11) and the change data regularization (CDR) of (12), we have the final GLGDE model as:

$$\min_{\mathbf{p} \in \mathbb{R}^{N_S}} \|\mathbf{p} - \bar{\mathbf{d}}\|_2^2 + \alpha \mathbf{p}^T \mathbf{L}_f \mathbf{p} + \beta \mathbf{p}^T \mathbf{L}_s \mathbf{p} \tag{13}$$

where $\alpha, \beta > 0$ are the balance parameters that control the weights of the GFGR and LSGR, respectively. Here, the larger the α and β, the smoother the DI on the graphs of G_f and G_s, respectively. The closed-form solution of the minimization problem (13) is:

$$\mathbf{p}^* = \left(\mathbf{I}_{N_S} + \alpha \mathbf{L}_f + \beta \mathbf{L}_s \right)^{-1} \bar{\mathbf{d}} \tag{14}$$

where $\mathbf{I}_{N_S}$ denotes the $N_S \times N_S$ identity matrix. Because p_i^* represents the change level of the i-th superpixel, we can obtain the enhanced DI as $\mathbf{D}^*$:

$$d^*(m, n) = p_i^*, \text{ if}(m, n) \in \Lambda_i \tag{15}$$

From the solution of $\mathbf{p}^*$ and $\mathbf{D}^*$, we can find that it is a process of improving the DI using the graph model, i.e., requiring the DI to be smooth on two different graphs of G_f and G_s.

Once the enhanced DI is computed, the binary CM can be calculated by using some image-segmentation methods, for example, the thresholding methods such as the OTSU threshold [48] or the kittler and Illingworth (KI) threshold [49], the clustering methods such as k-means clustering [50] or the Fuzzy c-means clustering [51], or the random field-based methods such as the Markov random field [37] or conditional random field [52]. Since the focus of this paper is on how to improve DI quality and not on how to segment DI, we directly use the OTSU method to classify the DI into changed and unchanged classes to obtain the final CM. The framework of the global and local graph-based DI-enhancement for CD is summarized in Algorithm 1.

Algorithm 1: GLGDE-based CD.

Input: Images of $\mathbf{X}$ and $\mathbf{Y}$, initial DI of $\mathbf{D}$.
      Parameters of N_S, α, and β.
Pre-processing:
    Segment $\mathbf{X}$, $\mathbf{Y}$, and $\mathbf{D}$ into superpixels with GMMSP.
    Extract the features to obtain $\tilde{\mathbf{X}}$ and $\tilde{\mathbf{Y}}$.
Graph construction:
    Find the KNN sets of $\mathcal{N}_i^x$ and $\mathcal{N}_i^y$.
    Find the R-adjacent neighbors of $\mathcal{N}_i^R$.
    Construct the graphs of G_f and G_s.
Model solving:
    Compute the $\mathbf{p}^*$ by using (14).
    Compute the $\mathbf{D}^*$ by using (15).
    Compute final CM by using OTSU thresholding method.

---

3. Experimental Results and Discussions

In this section, we demonstrate the performance of the proposed GLGDE through experiments, which are conducted on different CD tasks (homogeneous CD of SAR images and optical images and heterogeneous CD) with six datasets, as listed in Table 1.

Table 1. Description of the six datasets.

Dataset	Sensor (or Source)	Size (Pixels)	Date	Location	Event (and Spatial Resolution)
#1	Radarsat-2/Radarsat-2	$289 \times 257 \times 1(1)$	June 2008–June 2009	Yellow River, China	Farmland irrigation (8 m)
#2	Radarsat-2/Radarsat-2	$291 \times 306 \times 1(1)$	June 2008–June 2009	Yellow River, China	Farmland irrigation (8 m)
#3	Google Earth/Google Earth	$500 \times 500 \times 3(3)$	September 2012–March 2013	Beijing, China	Construction (1 m)
#4	Google Earth/Google Earth	$500 \times 500 \times 3(3)$	September 2012–March 2013	Beijing, China	Construction (1 m)
#5	Landsat-5/Google Earth	$300 \times 412 \times 1(3)$	September 1995–July 1996	Sardinia, Italy	Lake expansion (30 m)
#6	Radarsat-2/Google Earth	$593 \times 921 \times 1(3)$	June 2008–September 2012	Shuguang Village, China	Building construction (8 m)

3.1. Experimental Settings

The main parameters of the GLGDE are the number of superpixels N_S, the balance parameters of α and β. For all the experimental results, we set $N_S = 5000$, $\alpha = 0.5$, and $\beta = \frac{\sum_{(i,j) \in E_f} W_{i,j}^f}{\sum_{(i,j) \in E_s} W_{i,j}^s} \times \alpha$ for GLGDE. The impact of these parameters will be discussed in Section 3.3.

To measure the effect of DI enhancement, we use two types of evaluation metrics. First, we evaluate the quality of the DI directly by using: (1) the precision-recall (PR) curve along with the areas under the PR curve (AUP); (2) the empirical receiver operating characteristics (ROC) curve along with the areas under the ROC curve (AUR). Second, we evaluate the quality of the DI indirectly by assessing the CM obtained using the OTSU thresholding, which can be measured using the False alarm rate (Fa), Miss rate (Mr), overall accuracy (Oa), and the Kappa coefficient (Kc) computed using $\mathrm{Fa} = \mathrm{FP}/(\mathrm{FP} + \mathrm{TN})$, $\mathrm{Mr} = \mathrm{FN}/(\mathrm{TP} + \mathrm{FN})$, $\mathrm{Oa} = (\mathrm{TP} + \mathrm{TN})/(\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN})$, and $\mathrm{Kc} = (\mathrm{Oa} - \mathrm{PRE})/(1 - \mathrm{PRE})$ with:

$$\mathrm{PRE} = \frac{(\mathrm{TP} + \mathrm{FN})(\mathrm{TP} + \mathrm{FP}) + (\mathrm{TN} + \mathrm{FP})(\mathrm{TN} + \mathrm{FN})}{(\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN})^2}, \quad (16)$$

where TN, TP, FN, and FP represent the true negatives, true positives, false negatives, and false positives, respectively. At the same time, in the CM, we mark the TN, TP, FN, and FP with different colors.

3.2. Experimental Results

3.2.1. Homogeneous CD of SAR Images

Two pairs of SAR images are used in this task. Datasets #1 and #2 are both collected using a Radarsat-2 SAR sensor over the Yellow River Estuary, China, as shown in Figure 2. The pre-event images and post-event images are acquired in June 2008 and June 2009, respectively, the spatial resolution of the images is 8 m/pixel, and the ground change maps indicate the newly irrigated areas over Yellow River Estuary.

To obtain the initial DI, we choose the difference operator (Diff), log-ratio operator (LR), mean-ratio operator (MR) [26], neighborhood-ratio operator (NR) [53], sparsity-driven change-detection (SDCD) method [27], and the improved nonlocal patch-based graph (INLPG) model [11] as the baselines. In the MR and NR, the patch size is set to 3×3; in the SDCD, the regularization parameter λ manually selects the best examples by adjusting it from 10^{-4} to 10^{-1} with 20 logarithmic intervals; in the INLPG, the default parameters of the official code are used directly. After the initial DI is computed, the proposed GLGDE is applied to each DI to obtain the enhanced DI, denoted as the E-Diff, E-LR, E-MR, E-NR, E-SDCD, and E-INLPG, respectively.

Figure 3 shows the initial DIs and enhanced DIs of Datasets #1 and #2, Figure 4 shows the ROC and PR curves of these DIs, and Table 2 lists the corresponding AUR and AUP. From these results, we have three findings: first, the qualities of the enhanced DIs generated using the proposed GLGDE are much higher than that of the initial DIs, which means that the GLGDE can increase the contrast between changed and unchanged in the DI. For example, by comparing the DIs generated using Diff and E-Diff in Figure 3a, it can be seen that, with the former, it is difficult to detect changes, while the latter can clearly highlight the changes. Second, GLGDE is not only a local spatial smoothing of DI but, more importantly, it can correct the DI, that is, it can modify the DI by introducing the correlation

information of the original multi-temporal images, as illustrated by Figure 3e. Third, the poorer the initial DI performance, the more significant the improvement of the GLGDE. It can also be found that INLPG can achieve better performance in these initial DI-generation methods.
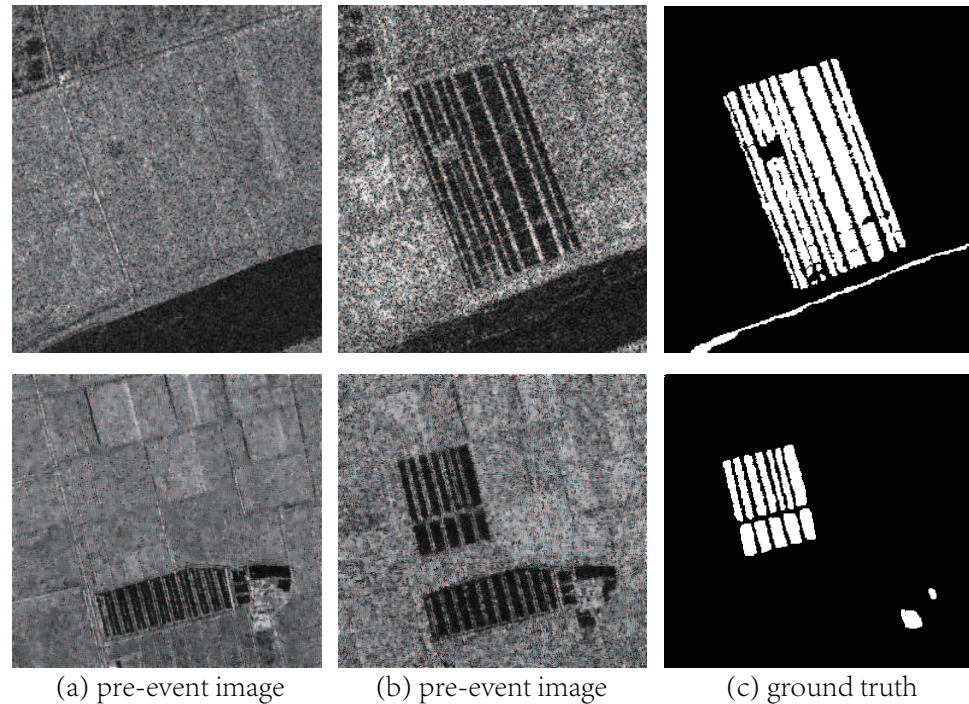

(a) pre-event image   (b) pre-event image   (c) ground truth

Figure 2. Datasets #1 (top row) and #2 (bottom row). From left to right are: (**a**) pre-event image; (**b**) post-event image; and (**c**) ground truth.

Table 2. Quantitative measures of DIs and CMs on the Datasets #1 and #2. Avg.ipv represents the average improvement.

Methods	Dataset #1						Dataset #2					
	AUR ↑	AUP ↑	Fa ↓	Mr ↓	Oa ↑	Kc ↑	AUR ↑	AUP ↑	Fa ↓	Mr ↓	Oa ↑	Kc ↑
Diff	0.657	0.248	0.317	0.453	0.659	0.167	0.818	0.194	0.326	0.158	0.684	0.154
LR	0.764	0.478	0.185	0.404	0.775	0.351	0.916	0.525	0.143	0.156	0.856	0.352
MR	0.902	0.805	0.226	0.143	0.789	0.470	0.966	0.844	0.263	0.038	0.750	0.238
NR	0.905	0.794	0.166	0.178	0.832	0.536	0.978	0.860	0.114	0.047	0.890	0.460
SDCD	0.900	0.601	0.236	0.098	0.789	0.484	0.971	0.788	0.188	0.033	0.821	0.326
INLPG	0.978	0.938	0.008	0.264	0.946	0.798	0.990	0.909	0.013	0.163	0.978	0.809
E-Diff	0.959	0.881	0.021	0.252	0.937	0.774	0.986	0.922	0.005	0.154	0.986	0.869
E-LR	0.971	0.911	0.017	0.228	0.945	0.802	0.993	0.943	0.005	0.172	0.985	0.863
E-MR	0.973	0.929	0.015	0.180	0.955	0.841	0.990	0.945	0.004	0.121	0.989	0.898
E-NR	0.981	0.938	0.012	0.183	0.957	0.847	0.994	0.957	0.003	0.135	0.989	0.897
E-SDCD	0.976	0.930	0.023	0.135	0.957	0.853	0.992	0.947	0.006	0.105	0.988	0.893
E-INLPG	0.983	0.945	0.013	0.182	0.956	0.845	0.996	0.962	0.004	0.147	0.988	0.885
Avg.ipv	0.123	0.278	−0.173	−0.063	0.153	0.359	0.052	0.259	−0.170	0.040	0.157	0.494

Figure 5 shows the binary CMs generated from the initial and enhanced DIs by using the OTSU thresholding on Datasets #1 and #2, where we mark the false positives (FP) and false negatives (FN) with different colors (red and cyan) for easy observation and comparison. Table 2 lists the corresponding Fa, Mr, Oa, and Kc. We can see that there are a large number of errors in the original CMs and, in the CMs obtained after using GLGDE enhancement, these errors are corrected and the resulting CMs are much more accurate. For

example, as reported in Table 2, the average improvements (Avg.ipv) of GLGDE in Dataset #1 on the AUR, AUP, Oa, and Kc metrics are 0.123, 0.278, 0.153, and 0.359, respectively.
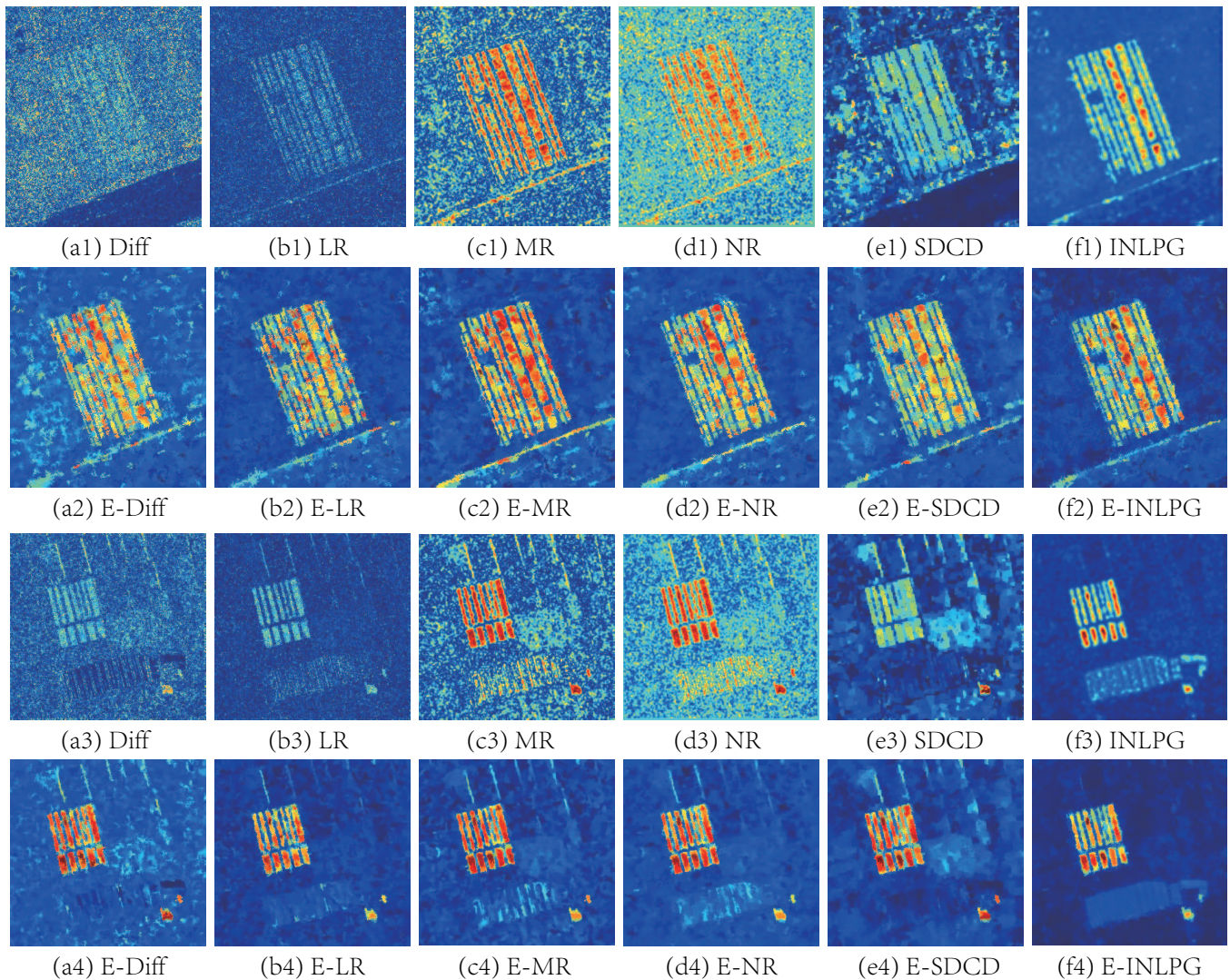

Figure 3. Initial and enhanced DIs of Datasets #1 and #2. From top to bottom, they correspond to initial DIs of Dataset #1, enhancement DIs of Dataset #1, initial DIs of Dataset #2, and enhancement DIs of Dataset #2. From left to right are DIs generated using: (**a1**–**a4**) Diff/E-Diff; (**b1**–**b4**) LR/E-LR; (**c1**–**c4**) MR/E-MR; (**d1**–**d4**) NR/E-NR; (**e1**–**e4**) SDCD/E-SDCD; and (**f1**–**f4**) INLPG/E-INLPG.
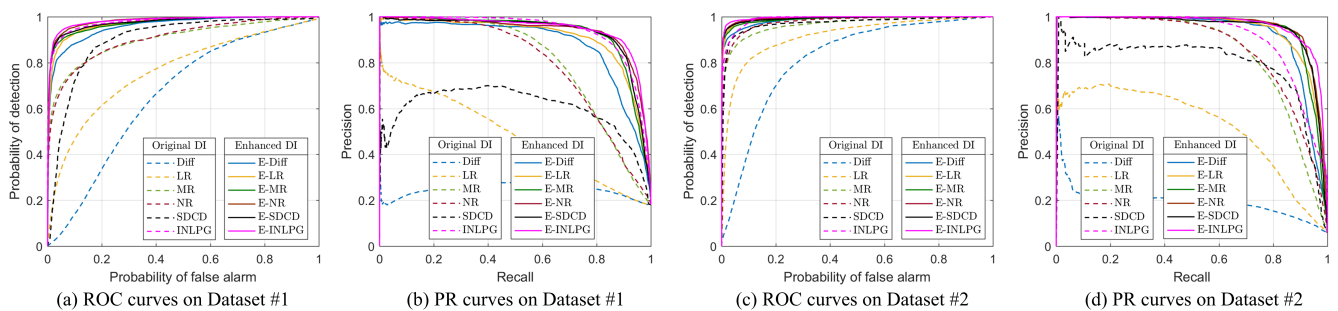
Figure 4. ROC and PR curves of Datasets #1 and #2. From left to right are: (**a**) ROC curves on Dataset #1; (**b**) PR curves on Dataset #1; (**c**) ROC curves on Dataset #2; and (**d**) PR curves on Dataset #2.
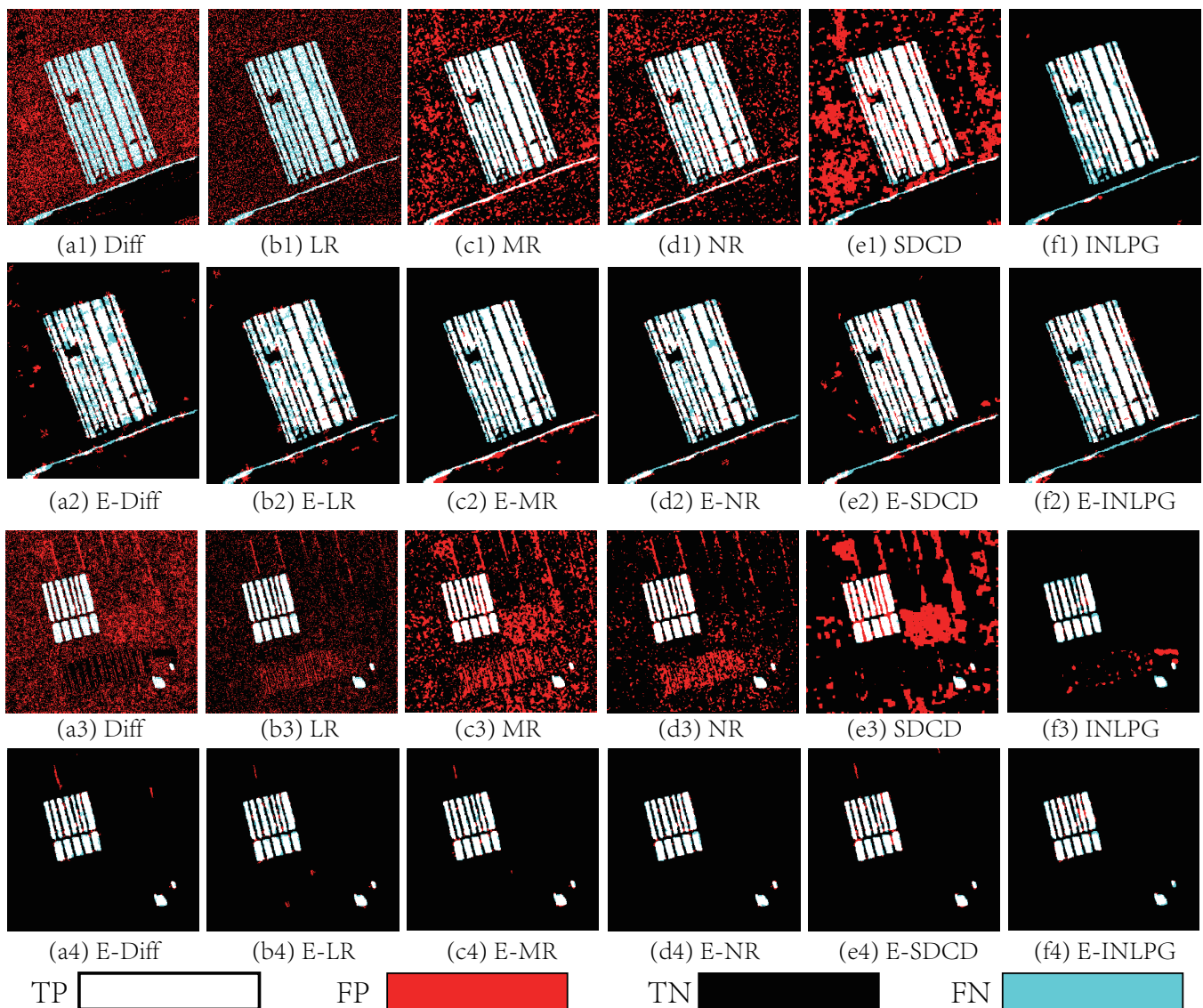
Figure 5. CMs computed from initial and enhanced DIs of Datasets #1 and #2. From top to bottom, they correspond to initial CMs of Dataset #1, enhancement CMs of Dataset #1, initial CMs of Dataset #2, and enhancement CMs of Dataset #2. From left to right are CMs generated by: (**a1–a4**) Diff/E-Diff; (**b1–b4**) LR/E-LR; (**c1–c4**) MR/E-MR; (**d1–d4**) NR/E-NR; (**e1–e4**) SDCD/E-SDCD; and (**f1–f4**) INLPG/E-INLPG. In the CM, White: true positives (TP); Red: false positives (FP); Black: true negatives (TN); and Cyan: false negatives (FN).

3.2.2. Homogeneous CD of Optical Images

Two pairs of optical images are used in this task. Datasets #3 and #4 are both collected from Google Earth over Beijing, China, as shown in Figure 6. The pre-event images and post-event images were acquired in September 2012 and March 2013, respectively, the spatial resolution of the images is 1m/pixel, and the ground change maps indicate the newly constructed buildings.
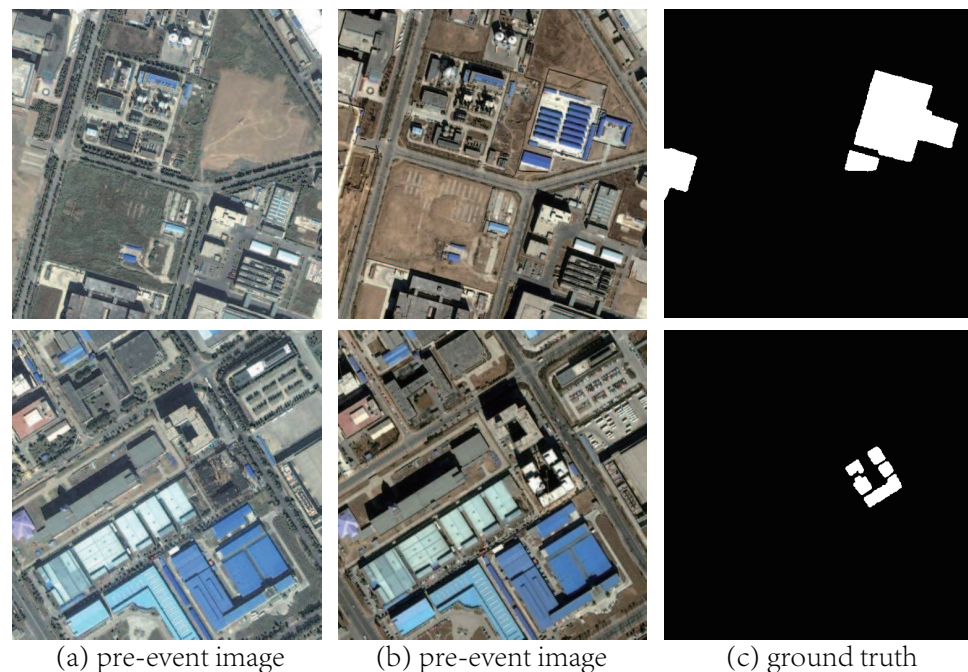
| (a) pre-event image | (b) pre-event image | (c) ground truth |

Figure 6. Datasets #3 (top row) and #4 (bottom row). From left to right are: (**a**) pre-event image; (**b**) post-event image; and (**c**) ground truth.
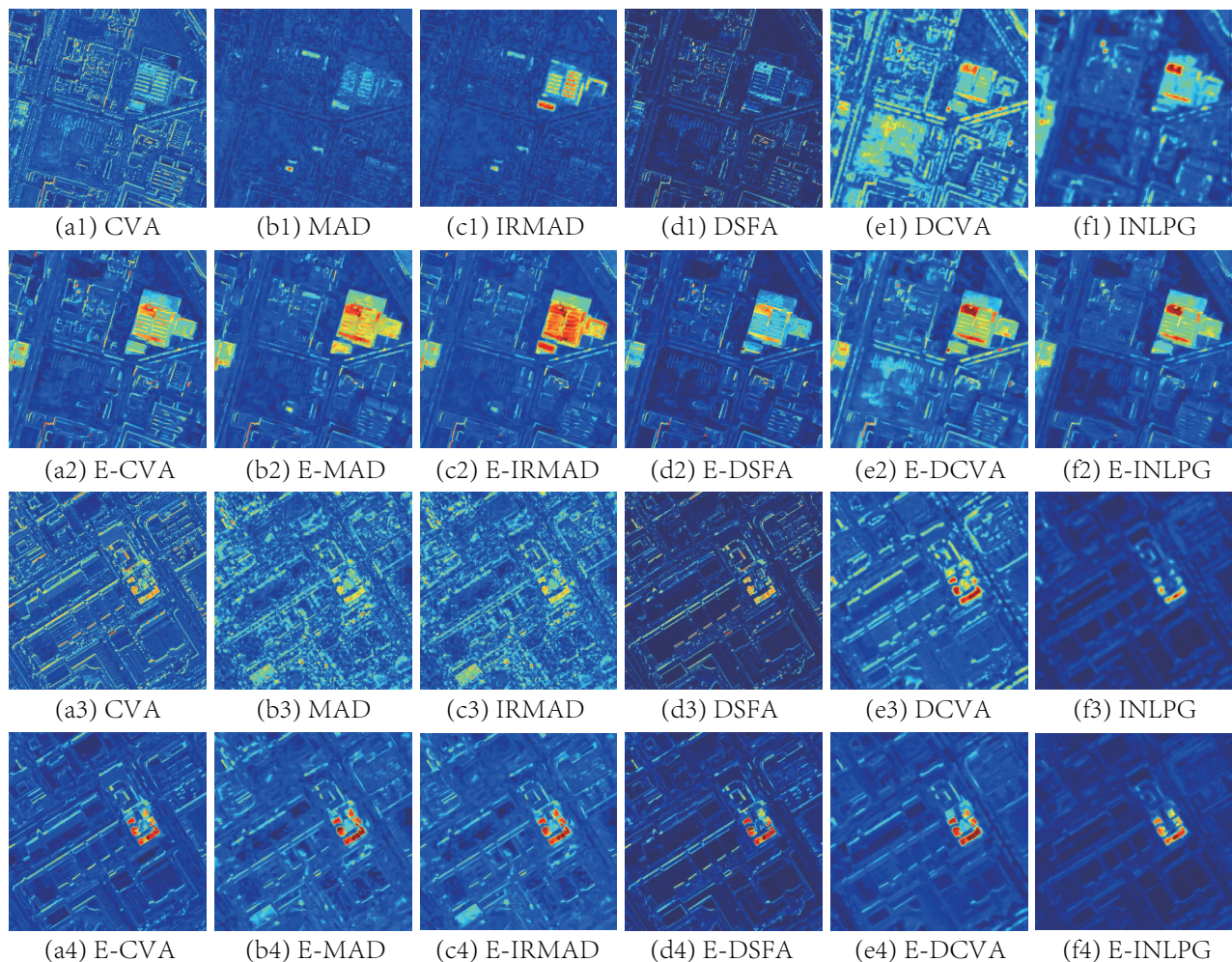
To obtain the initial DI, we choose the change vector analysis (CVA) [18], multivariate alteration detection method (MAD) [19], iteratively reweighted MAD method (IRMAD) [20], deep slow feature analysis network (DSFA) [23], deep CVA network (DCVA) [21], and the INLPG [11] as the baselines. After the initial DI is computed, the proposed GLGDE is applied to each DI to obtain the enhanced DI, denoted as the E-CVA, E-MAR, E-IRMAD, E-DSFA, E-DCVA, and E-INLPG, respectively.

Figure 7 shows the initial DIs and enhanced DIs of Datasets #3 and #4, Figure 8 shows the ROC and PR curves of these DIs, and Table 3 lists the corresponding AUR and AUP. As can be seen from Figures 6 and 7, the seasonal differences in the pre-event and post-event images lead to the poor qualities of the initial DIs obtained by some methods, such as the CVA and DCVA on Dataset #3 and MAD and IRMAD on Dataset #4. By using the proposed GLGDE on these initial DIs, the distinguishability between changed and unchanged is greatly improved, as shown in Figures 7 and 8 and especially illustrated by the PR curves of Figure 8. The average improvements of GLGDE in Datasets #3 and Datasets #4 on the AUP metric are 0.421 and 0.450, respectively.

Figure 9 shows the binary CMs of Datasets #3 and #4 generated by using the OTSU thresholding on the DIs of Figure 7 and Table 3 lists the corresponding Fa, Mr, Oa, and Kc. For Dataset #3, there are lots of false alarms in the initial CMs of CVA, MAD, and DCVA and lots of miss detection in the initial CMs of IRMAD and DSFA, though most of these errors are corrected after the enhancement of GLGDE. Therefore, it can be seen that GLGDE is not only a smoothing of DI but also a correction of DI. It enables the entire error area to be corrected, which is not possible in the common spatial smoothing methods that only use contextual information of a DI, as shown in Figures 7 and 9. The average improvements of GLGDE in Dataset #3 and Dataset #4 on the Kc metric are 0.414 and 0.148, respectively.

Table 3. Quantitative measures of DIs and CMs on the Datasets #3 and #4. Avg.ipv represents the average improvement.

Methods	Dataset #3						Dataset #4					
	AUR ↑	AUP ↑	Fa ↓	Mr ↓	Oa ↑	Kc ↑	AUR ↑	AUP ↑	Fa ↓	Mr ↓	Oa ↑	Kc ↑
CVA	0.712	0.160	0.191	0.474	0.787	0.185	0.798	0.127	0.168	0.332	0.830	0.073
MAD	0.885	0.420	0.193	0.198	0.806	0.312	0.859	0.192	0.317	0.166	0.685	0.042
IRMAD	0.911	0.687	0.006	0.569	0.950	0.549	0.856	0.224	0.306	0.179	0.696	0.043
DSFA	0.769	0.194	0.081	0.685	0.871	0.207	0.824	0.180	0.082	0.395	0.913	0.139
DCVA	0.715	0.212	0.335	0.329	0.666	0.127	0.955	0.614	0.140	0.095	0.860	0.127
INLPG	0.955	0.652	0.124	0.086	0.879	0.484	0.992	0.796	0.007	0.228	0.990	0.674
E-CVA	0.975	0.685	0.076	0.047	0.926	0.631	0.978	0.706	0.093	0.066	0.907	0.195
E-MAD	0.993	0.902	0.033	0.016	0.969	0.815	0.982	0.809	0.137	0.028	0.865	0.142
E-IRMAD	0.995	0.942	0.022	0.035	0.977	0.858	0.980	0.799	0.140	0.033	0.861	0.137
E-DSFA	0.976	0.680	0.073	0.018	0.932	0.658	0.979	0.702	0.047	0.160	0.952	0.306
E-DCVA	0.975	0.744	0.093	0.049	0.911	0.580	0.994	0.867	0.050	0.036	0.950	0.328
E-INLPG	0.992	0.898	0.035	0.012	0.966	0.804	0.999	0.949	0.001	0.194	0.997	0.876
Avg.ipv	0.160	0.421	−0.100	−0.361	0.120	0.414	0.105	0.450	−0.092	−0.146	0.093	0.148

(a1) CVA (b1) MAD (c1) IRMAD (d1) DSFA (e1) DCVA (f1) INLPG

(a2) E-CVA (b2) E-MAD (c2) E-IRMAD (d2) E-DSFA (e2) E-DCVA (f2) E-INLPG

(a3) CVA (b3) MAD (c3) IRMAD (d3) DSFA (e3) DCVA (f3) INLPG

(a4) E-CVA (b4) E-MAD (c4) E-IRMAD (d4) E-DSFA (e4) E-DCVA (f4) E-INLPG

Figure 7. Initial and enhanced DIs of Datasets #3 and #4. From top to bottom, they correspond to initial DIs of Dataset #3, enhancement DIs of Dataset #3, initial DIs of Dataset #4, and enhancement DIs of Dataset #4. From left to right are DIs generated using: (**a1–a4**) CVA/E-CVA; (**b1–b4**) MAD/E-MAD; (**c1–c4**) IRMAD/E-IRMAD; (**d1–d4**) DSFA/E-DSFA; (**e1–e4**) DCVA/E-DCVA; and (**f1–f4**) INLPG/E-INLPG.

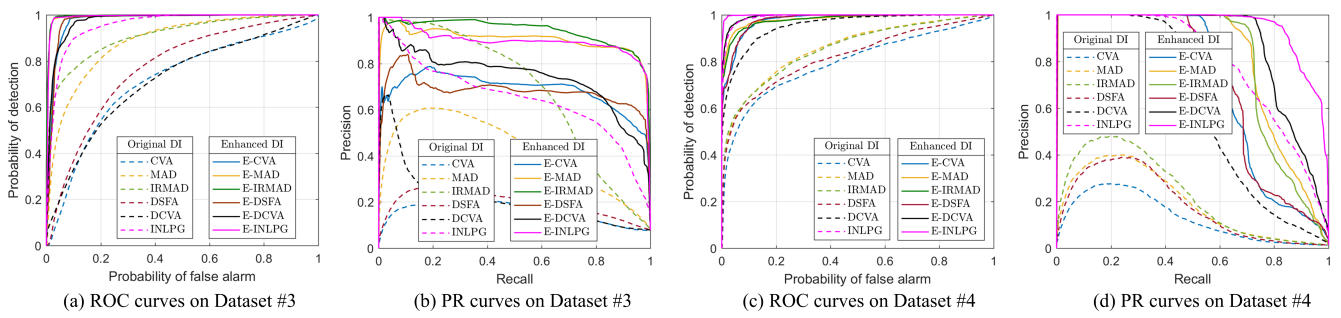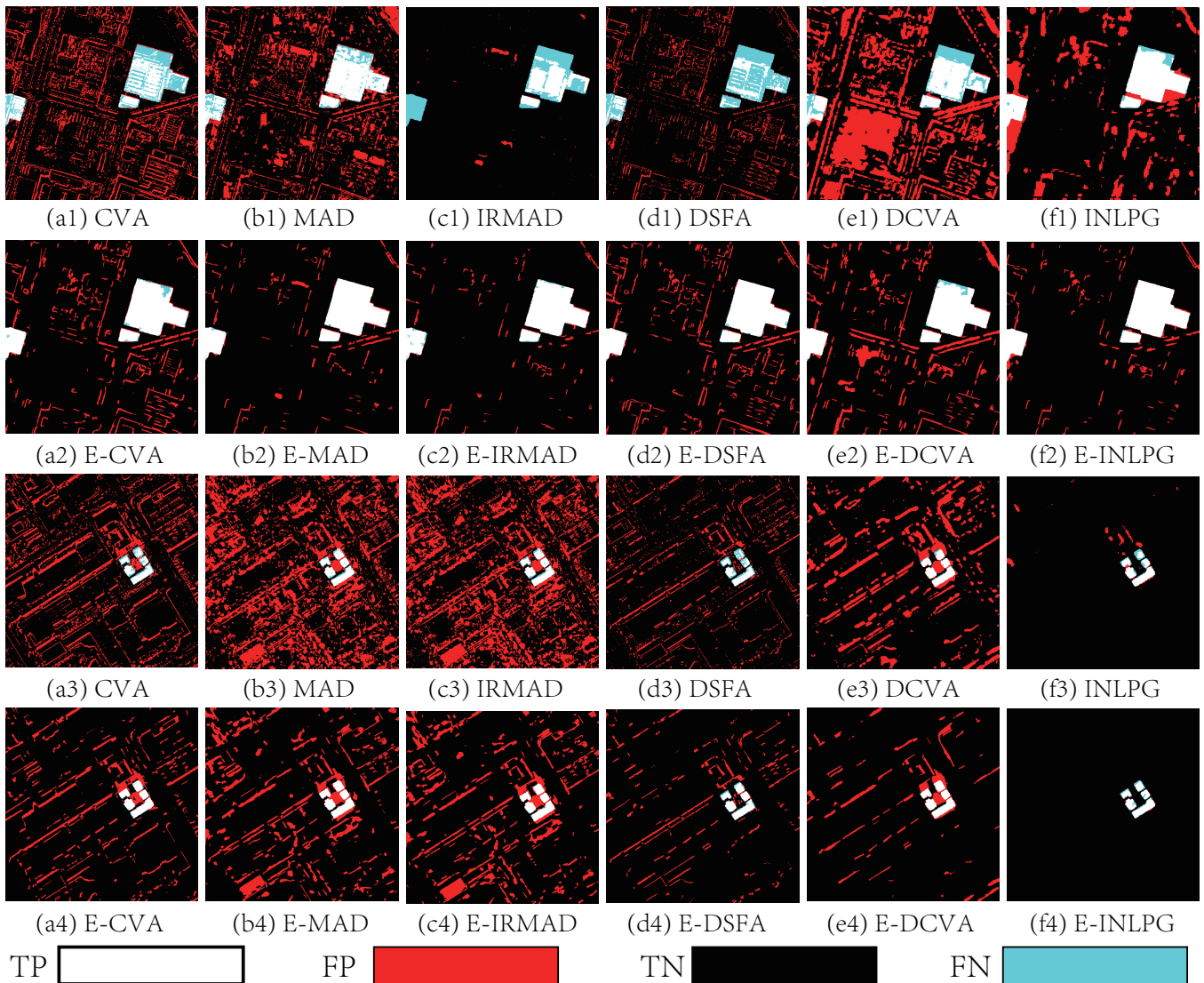| (a) ROC curves on Dataset #3 | (b) PR curves on Dataset #3 | (c) ROC curves on Dataset #4 | (d) PR curves on Dataset #4 |

Figure 8. ROC and PR curves of Datasets #3 and #4. From left to right are: (**a**) ROC curves on Dataset #3; (**b**) PR curves on Dataset #3; (**c**) ROC curves on Dataset #4; (**d**) PR curves on Dataset #4.

(a1) CVA	(b1) MAD	(c1) IRMAD	(d1) DSFA	(e1) DCVA	(f1) INLPG
(a2) E-CVA	(b2) E-MAD	(c2) E-IRMAD	(d2) E-DSFA	(e2) E-DCVA	(f2) E-INLPG
(a3) CVA	(b3) MAD	(c3) IRMAD	(d3) DSFA	(e3) DCVA	(f3) INLPG
(a4) E-CVA	(b4) E-MAD	(c4) E-IRMAD	(d4) E-DSFA	(e4) E-DCVA	(f4) E-INLPG

TP   FP   TN   FN

Figure 9. CMs computed from initial and enhanced DIs of Datasets #3 and #4. From top to bottom, they correspond to initial CMs of Dataset #3, enhancement CMs of Dataset #3, initial CMs of Dataset #4, enhancement CMs of Dataset #4. From left to right are CMs generated by: (**a1**–**a4**) CVA/E-CVA; (**b1**–**b4**) MAD/E-MAD; (**c1**–**c4**) IRMAD/E-IRMAD; (**d1**–**d4**) DSFA/E-DSFA; (**e1**–**e4**) DCVA/E-DCVA; and (**f1**–**f4**) INLPG/E-INLPG.

3.2.3. Heterogeneous CD

Two pairs of heterogeneous images are used in this task, as shown in Figure 10. In Dataset #5, the pre-event image is collected using Landsat-5 with the near-infrared band in September 1995, the post-event image is obtained from Google Earth using R, G, and B bands in July 1996, the spatial resolution of the images is 30 m/pixel, and the ground change map indicates the Lake expansion in Sardinia, Italy. In Dataset #6, the pre-event image is collected using a Radarsat-2 SAR sensor in June 2008, the post-event image is obtained from Google Earth with R, G, and B bands in September 2012, the spatial resolution of the images is 8 m/pixel, and the ground change map indicates the building construction in Shuguang Village, China.

(a) pre-event image   (b) pre-event image   (c) ground truth
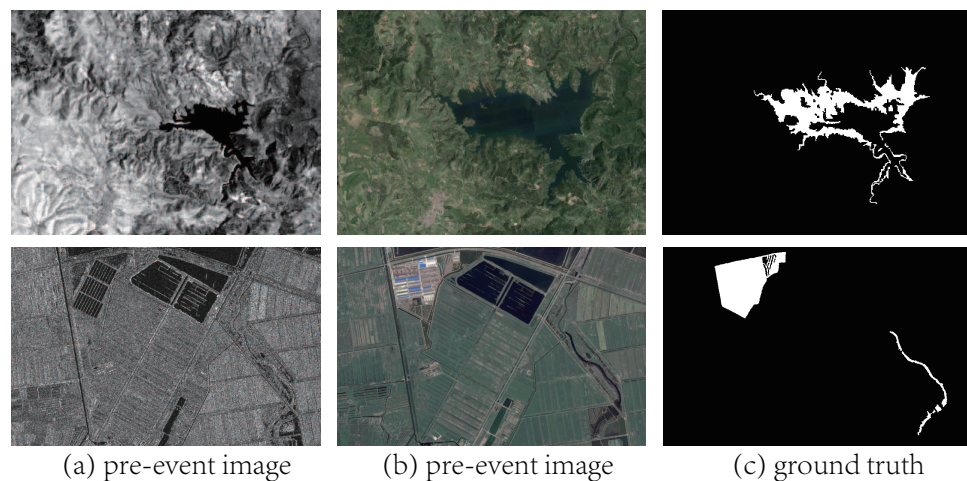
Figure 10. Datasets #5 (top row) and #6 (bottom row). From left to right are: (**a**) pre-event image; (**b**) post-event image; and (**c**) ground truth.

To obtain the initial DI, we choose the homogeneous pixel transformation (HPT) method [13], affinity matrix distance-based image regression (AMDIR) [14], adaptive local structure consistency-based method (ALSC) [54], INLPG [11], fractal projection and Markovian segmentation-based algorithm (FPMS) [34], and sparse-constrained adaptive structure consistency-based method (SCASC) [37] as the baselines. For the HPT, we use 40% of the unchanged pixels as the training samples; for other methods, we directly use the default parameters of their official codes. After the initial DI is computed, the proposed GLGDE is applied to each DI to obtain the enhanced DI, denoted as the E-HPT, E-AMDIR, E-ALSC, E-INLPG, E-FPMS, and E-SCASC, respectively.

Figure 11 shows the initial DIs and enhanced DIs of Datasets #5 and #6, Figure 12 shows the ROC and PR curves of these DIs, and Table 4 lists the corresponding AUR and AUP. From these results, we can see that the initial DIs all show some change information and, among these DIs, FPMS and INLPG perform relatively better. By comparing the initial DIs and enhanced DIs of Figure 11, two findings can be noted: first, the enhanced DIs generated using GLGDE can cause the change area to be more continuous than the corresponding initial DIs, such as the DIs obtained using SCASC and E-SCASC in Figure 11(f1–f4); second, the GLGDE can effectively suppress the unchanged regions with high change levels in the initial DI, thus reducing the interference of the background area on change detection, such as the DIs obtained using AMDIR and E-AMDIR in Figure 11(b1–b4). The ROC and PR curves in Figure 12 also verify the enhancement of GLGDE on DI and the average improvements of GLGDE in Datasets #5 and Datasets #6 on the AUP metric are 0.276 and 0.236, respectively.
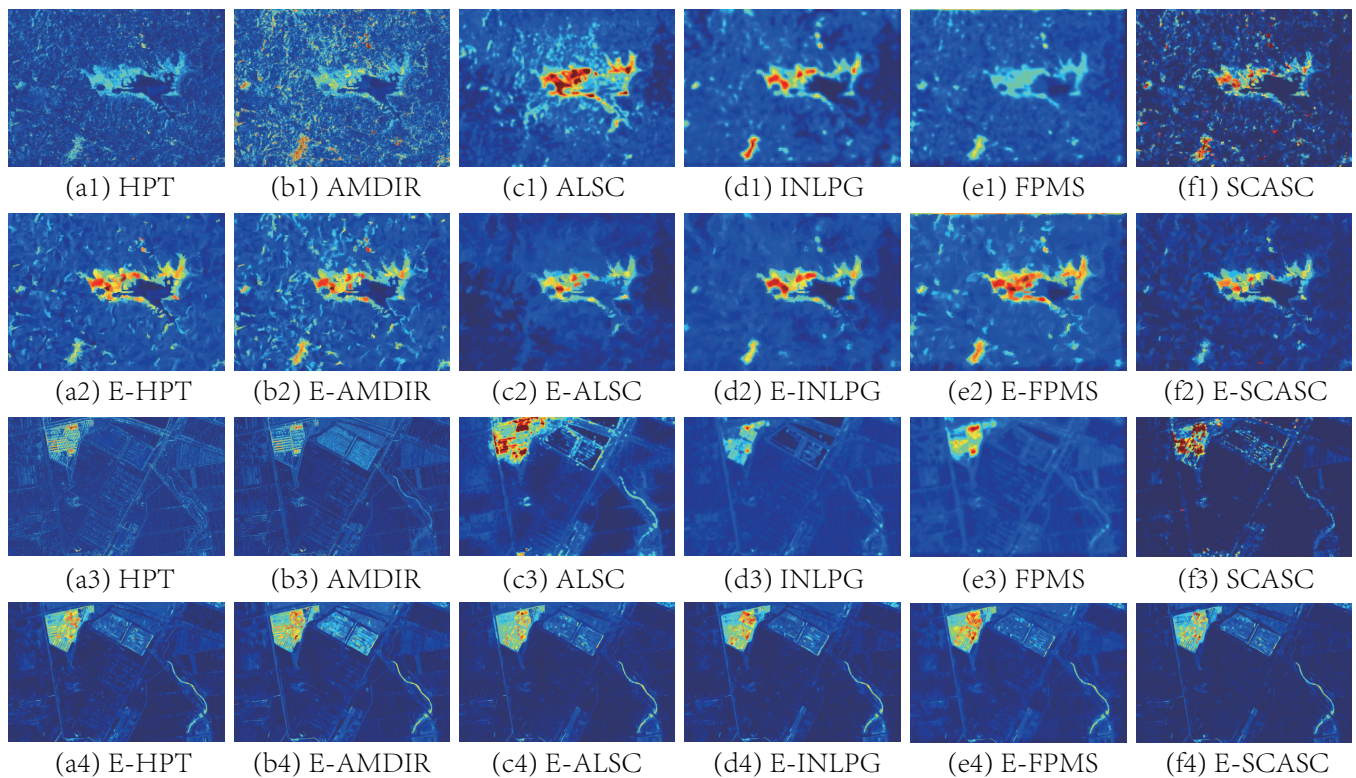
Figure 11. Initial and enhanced DIs of Datasets #5 and #6. From top to bottom, they correspond to initial DIs of Dataset #5, enhancement DIs of Dataset #5, initial DIs of Dataset #6, and enhancement DIs of Dataset #6. From left to right are DIs generated using: (**a1–a4**) CVA/E-CVA; (**b1–b4**) MAD/E-MAD; (**c1–c4**) IRMAD/E-IRMAD; (**d1–d4**) DSFA/E-DSFA; (**e1–e4**) DCVA/E-DCVA; and (**f1–f4**) INLPG/E-INLPG.

Figure 13 shows the CMs of Datasets #5 and #6 using different DIs with OTSU thresholding and Table 4 lists the corresponding quantitative measures. It can be seen that the GLGDE can effectively improve the detection performance. For example, GLGDE can significantly reduce false alarms in HPT and AMDIR in Figures 13(a1–a4) and 13(b1–b4), respectively, and preserve the edges of the detected area, such as the results on Dataset #6. As reported in Table 4, the average improvements of GLGDE in Datasets #5 and #6 on the Kc metric are 0.196 and 0.252, respectively.
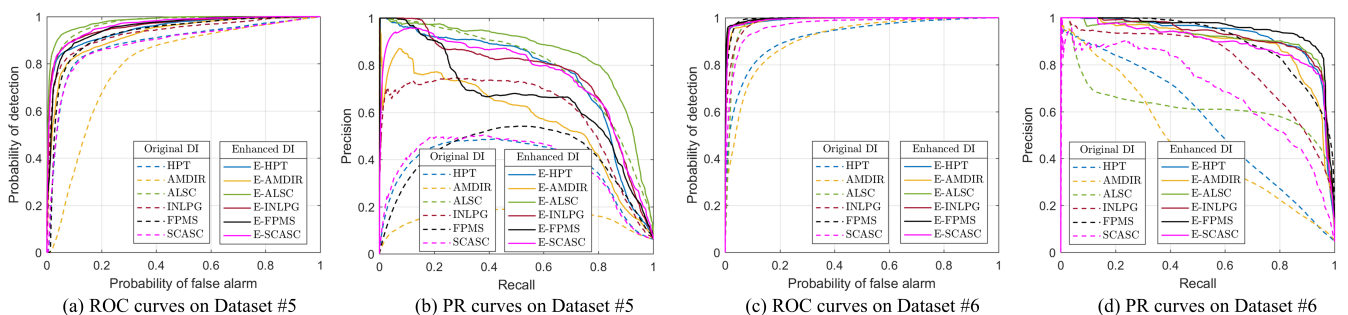

Figure 12. ROC and PR curves of Datasets #5 and #6. From left to right are: (**a**) ROC curves on Dataset #5; (**b**) PR curves on Dataset #5; (**c**) ROC curves on Dataset #6; and (**d**) PR curves on Dataset #6.
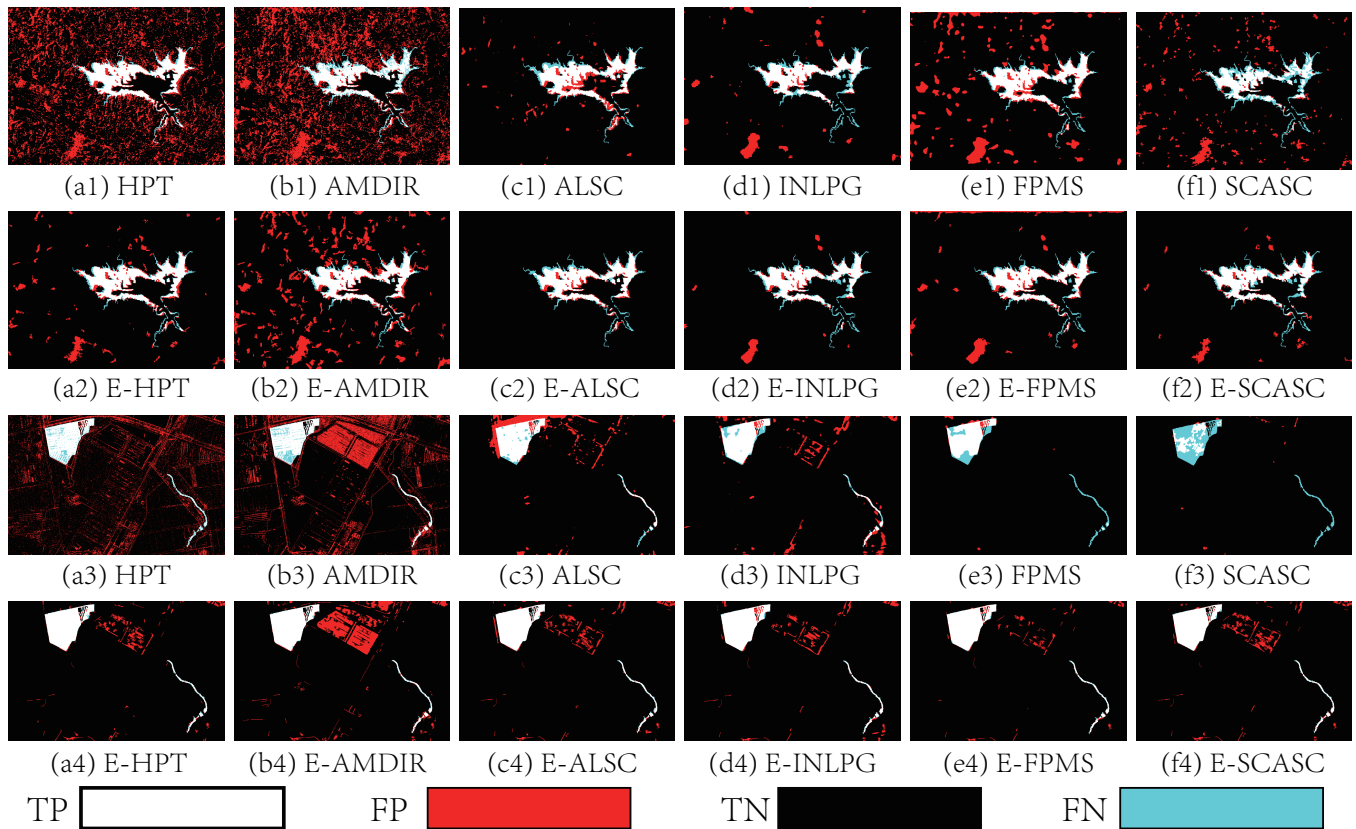
Figure 13. CMs computed from initial and enhanced DIs of Datasets #5 and #6. From top to bottom, they correspond to initial CMs of Dataset #5, enhancement CMs of Dataset #5, initial CMs of Dataset #6, and enhancement CMs of Dataset #6. From left to right are CMs generated by: (**a1–a4**) CVA/E-CVA; (**b1–b4**) MAD/E-MAD; (**c1–c4**) IRMAD/E-IRMAD; (**d1–d4**) DSFA/E-DSFA; (**e1–e4**) DCVA/E-DCVA; and (**f1–f4**) INLPG/E-INLPG.

Table 4. Quantitative measures of DIs and CMs on the Datasets #5 and #6. Avg.ipv represents the average improvement.

Methods	Dataset #5						Dataset #6					
	AUR ↑	AUP ↑	Fa ↓	Mr ↓	Oa ↑	Kc ↑	AUR ↑	AUR↑	Fa ↓	Mr ↓	Oa ↑	Kc ↑
HPT	0.889	0.373	0.196	0.138	0.808	0.286	0.922	0.564	0.117	0.181	0.880	0.339
AMDIR	0.795	0.155	0.224	0.277	0.773	0.203	0.911	0.470	0.151	0.180	0.848	0.279
ALSC	0.972	0.793	0.026	0.218	0.962	0.696	0.980	0.623	0.038	0.115	0.958	0.641
INLPG	0.930	0.604	0.027	0.297	0.956	0.642	0.985	0.808	0.029	0.151	0.965	0.672
FPMS	0.925	0.406	0.083	0.189	0.911	0.486	0.994	0.904	0.004	0.310	0.982	0.768
SCASC	0.885	0.383	0.043	0.420	0.934	0.485	0.968	0.695	0.002	0.713	0.965	0.418
E-HPT	0.948	0.757	0.035	0.186	0.956	0.672	0.992	0.922	0.019	0.044	0.979	0.800
E-AMDIR	0.926	0.576	0.092	0.181	0.903	0.465	0.991	0.889	0.060	0.039	0.941	0.571
E-ALSC	0.980	0.851	0.011	0.229	0.975	0.781	0.994	0.911	0.021	0.040	0.979	0.794
E-INLPG	0.959	0.761	0.017	0.261	0.968	0.723	0.994	0.920	0.020	0.038	0.980	0.802
E-FPMS	0.953	0.669	0.040	0.221	0.949	0.626	0.996	0.948	0.011	0.039	0.987	0.869
E-SCASC	0.964	0.755	0.018	0.280	0.966	0.705	0.992	0.892	0.021	0.038	0.978	0.792
Avg.ipv	0.056	0.276	−0.064	−0.030	0.062	0.196	0.033	0.236	−0.032	−0.235	0.041	0.252

3.3. Parameter Analysis

The main parameters of the proposed GLGDE are the number of superpixels N_S and the balance parameters of α and β.

Generally, the N_S should be selected according to the image resolution and granularity requirement of the CD task. A larger N_S cause the segmented superpixel to be smaller,

which improves the detection granularity but also increases the computational complexity. In addition, we have the following notes for the influence of the image resolution on the proposed method. First, high-resolution images usually contain more detailed structure and textural information, so when performing superpixel co-segmentation in the pre-processing, the segmented superpixels are required to be finer, i.e., N_S is larger. Second, when the spatial resolution of the compared image is very high, the great intraclass variation and low interclass variance usually lead to a lot of salt-and-pepper noise in the original DI. At the same time, it is difficult to accurately characterize the structure of high-resolution images with the KNN graphs. Therefore, when facing the very high-resolution images, the following two issues need to be considered in the proposed method: first, the number of superpixels N_S should be increased appropriately; second, the KNN graphs used to capture the structures of multi-temporal images and the DI may need to be replaced with the more advanced graph neural networks (GNN).

The balance parameters of α and β are used to control the weights of the GFGR and LSGR in the DI-enhancement model (13), respectively. First, in order to cause the GFGR and LSGR to be equivalent (the penalty is balanced), we set $\beta = \frac{\sum_{(i,j)\in E_f} W_{i,j}^f}{\sum_{(i,j)\in E_s} W_{i,j}^s} \times \alpha$, which makes causes the values of $\alpha \mathbf{p}^T \mathbf{L}_f \mathbf{p}$ and $\beta \mathbf{p}^T \mathbf{L}_s \mathbf{p}$ to be approximately at the same level. Second, to measure the impact of α, we adjust the α from 2^{-4} to 2^2 with ratio of 2 and plot the average improvement of the AUP generated using GLGDE with different α in Figure 14. It can be found that too large and too small α are both not suitable: first, if α is too large, the GLGDE tends to over-smooth the DI, causing all the superpixels connected by the constructed graphs (G_f and G_s) to converge to the same; second, if α is too small, then GFGR and LSGR cannot play a noticeable effect in the DI-enhancement model, which limits the performance of GLGDE. Based on Figure 14, we fix $\alpha = 0.5$ in this paper for simplicity.
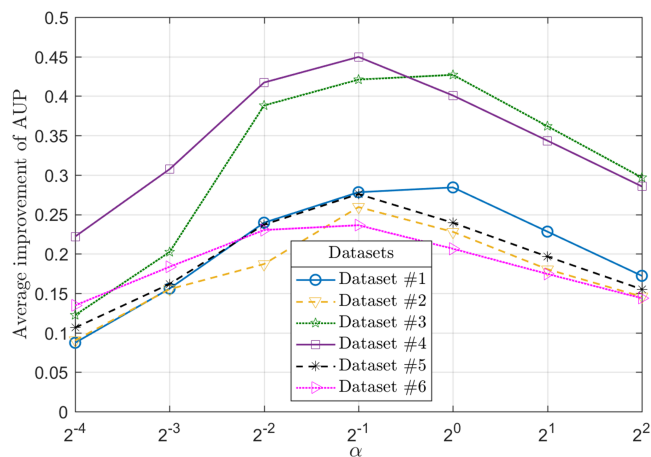

Figure 14. Sensitivity analysis of parameter α in GLGDE.

4. Conclusions

In this paper, we focus on the post-processing of the DI that is rarely noticed by other methods and propose a DI-enhancement method, named GLGDE. Once the initial DI is obtained, GLGDE first segments the DI and multi-temporal images into superpixels and then constructs a global feature graph and a local spatial graph with superpixels as vertices for the DI, which can exploit the change information and contextual information in the DI and the correlation information in the multi-temporal images. With the constructed graphs, the DI-enhancement model is built using three terms: a global-feature-graph-induced regularization term, a local spatial-graph-induced regularization term, and a change-data regularization term. By solving the minimization model, we can obtain the improved DI. Different from the previous DI-smoothing algorithm that only uses the contextual informa-tion of the DI itself, the proposed GLGDE can also exploit the association information of the

multi-temporal images, so it can not only smooth the DI but also correct the DI. Therefore, it is a DI-enhancement method specifically designed for a CD problem, which takes into account the characteristics of the CD problem. Finally, extensive experiments in different CD scenarios demonstrate the effectiveness of the proposed method. In our future study, we will exploit the global feature graph and local spatial graph in the DI segmentation process of a CD problem to improve the detection performance of the final change map.

Author Contributions: Methodology, X.Z.; software, X.Z. and D.G.; validation, X.Z., B.L. and Z.C. Original draft preparation, X.Z. and D.G.; writing—review and editing, L.P. and D.G.; supervision, B.L. and L.P. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

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

References

1. Singh, A. Review Article Digital change detection techniques using remotely-sensed data. *Int. J. Remote Sens.* **1989**, *10*, 989–1003. [CrossRef]
2. Bovolo, F.; Bruzzone, L. The time variable in data fusion: A change detection perspective. *IEEE Geosci. Remote Sens. Mag.* **2015**, *3*, 8–26. [CrossRef]
3. Kennedy, R.E.; Townsend, P.A.; Gross, J.E.; Cohen, W.B.; Bolstad, P.; Wang, Y.; Adams, P. Remote sensing change detection tools for natural resource managers: Understanding concepts and tradeoffs in the design of landscape monitoring projects. *Remote Sens. Environ.* **2009**, *113*, 1382–1396. [CrossRef]
4. Gil-Yepes, J.L.; Ruiz, L.A.; Recio, J.A.; Balaguer-Beser, Á.; Hermosilla, T. Description and validation of a new set of object-based temporal geostatistical features for land-use/land-cover change detection. *ISPRS J. Photogramm. Remote Sens.* **2016**, *121*, 77–91. [CrossRef]
5. Taubenböck, H.; Esch, T.; Felbier, A.; Wiesner, M.; Roth, A.; Dech, S. Monitoring urbanization in mega cities from space. *Remote Sens. Environ.* **2012**, *117*, 162–176. [CrossRef]
6. Zhang, P.; Ban, Y.; Nascetti, A. Learning U-Net without forgetting for near real-time wildfire monitoring by the fusion of SAR and optical time series. *Remote Sens. Environ.* **2021**, *261*, 112467. [CrossRef]
7. Lv, Z.; Liu, T.; Benediktsson, J.A.; Falco, N. Land Cover Change Detection Techniques: Very-high-resolution optical images: A review. *IEEE Geosci. Remote Sens. Mag.* **2022**, *10*, 44–63. [CrossRef]
8. You, Y.; Cao, J.; Zhou, W. A survey of change detection methods based on remote sensing images for multi-source and multi-objective scenarios. *Remote Sens.* **2020**, *12*, 2460. [CrossRef]
9. Shafique, A.; Cao, G.; Khan, Z.; Asad, M.; Aslam, M. Deep learning-based change detection in remote sensing images: A review. *Remote Sens.* **2022**, *14*, 871. [CrossRef]
10. 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]
11. Sun, Y.; Lei, L.; Li, X.; Tan, X.; Kuang, G. Structure Consistency-Based Graph for Unsupervised Change Detection with Homogeneous and Heterogeneous Remote Sensing Images. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 4700221. [CrossRef]
12. Bruzzone, L.; Prieto, D. An adaptive semiparametric and context-based approach to unsupervised change detection in multitemporal remote-sensing images. *IEEE Trans. Image Process.* **2002**, *11*, 452–466. [CrossRef]
13. Liu, Z.; Li, G.; Mercier, G.; He, Y.; Pan, Q. Change Detection in Heterogenous Remote Sensing Images via Homogeneous Pixel Transformation. *IEEE Trans. Image Process.* **2018**, *27*, 1822–1834. [CrossRef]
14. Luppino, L.T.; Bianchi, F.M.; Moser, G.; Anfinsen, S.N. Unsupervised Image Regression for Heterogeneous Change Detection. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 9960–9975. [CrossRef]
15. Lv, Z.; Wang, F.; Liu, T.; Kong, X.; Benediktsson, J.A. Novel Automatic Approach for Land Cover Change Detection by Using VHR Remote Sensing Images. *IEEE Geosci. Remote Sens. Lett.* **2022**, *19*, 1–5. [CrossRef]
16. Luppino, L.T.; Hansen, M.A.; Kampffmeyer, M.; Bianchi, F.M.; Moser, G.; Jenssen, R.; Anfinsen, S.N. Code-Aligned Autoencoders for Unsupervised Change Detection in Multimodal Remote Sensing Images. *IEEE Trans. Neural Netw. Learn. Syst.* **2022**, 1–13. [CrossRef] [PubMed]
17. Luppino, L.T.; Kampffmeyer, M.; Bianchi, F.M.; Moser, G.; Serpico, S.B.; Jenssen, R.; Anfinsen, S.N. Deep Image Translation with an Affinity-Based Change Prior for Unsupervised Multimodal Change Detection. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 4700422. [CrossRef]

18. Malila, W.A. Change vector analysis: An approach for detecting forest changes with Landsat. In Proceedings of the LARS Symposia, West Lafayette, IN, USA, 3–6 June 1980, p. 385.

19. Nielsen, A.A.; Conradsen, K.; Simpson, J.J. Multivariate alteration detection (MAD) and MAF postprocessing in multispectral, bitemporal image data: New approaches to change detection studies. *Remote Sens. Environ.* **1998**, *64*, 1–19. [CrossRef]

20. Nielsen, A.A. The Regularized Iteratively Reweighted MAD Method for Change Detection in Multi- and Hyperspectral Data. *IEEE Trans. Image Process.* **2007**, *16*, 463–478. [CrossRef]

21. Saha, S.; Bovolo, F.; Bruzzone, L. Unsupervised Deep Change Vector Analysis for Multiple-Change Detection in VHR Images. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 3677–3693. [CrossRef]

22. 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]

23. Du, B.; Ru, L.; Wu, C.; Zhang, L. Unsupervised Deep Slow Feature Analysis for Change Detection in Multi-Temporal Remote Sensing Images. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 9976–9992. [CrossRef]

24. Moser, G.; Serpico, S. Generalized minimum-error thresholding for unsupervised change detection from SAR amplitude imagery. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 2972–2982. [CrossRef]

25. Bovolo, F.; Bruzzone, L. A detail-preserving scale-driven approach to change detection in multitemporal SAR images. *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*, 2963–2972. [CrossRef]

26. Inglada, J.; Mercier, G. A New Statistical Similarity Measure for Change Detection in Multitemporal SAR Images and Its Extension to Multiscale Change Analysis. *IEEE Trans. Geosci. Remote Sens.* **2007**, *45*, 1432–1445. [CrossRef]

27. Nar, F.; Özgür, A.; Saran, A.N. Sparsity-Driven Change Detection in Multitemporal SAR Images. *IEEE Geosci. Remote Sens. Lett.* **2016**, *13*, 1032–1036. [CrossRef]

28. Sun, Y.; Lei, L.; Guan, D.; Li, X.; Kuang, G. SAR Image Change Detection Based on Nonlocal Low-Rank Model and Two-Level Clustering. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *13*, 293–306. [CrossRef]

29. Zhang, W.; Jiao, L.; Liu, F.; Yang, S.; Liu, J. Adaptive Contourlet Fusion Clustering for SAR Image Change Detection. *IEEE Trans. Image Process.* **2022**, *31*, 2295–2308. [CrossRef] [PubMed]

30. Wang, J.; Zhao, T.; Jiang, X.; Lan, K. A Hierarchical Heterogeneous Graph for Unsupervised SAR Image Change Detection. *IEEE Geosci. Remote. Sens. Lett.* **2022**, *19*, 4516605. [CrossRef]

31. Planinšič, P.; Gleich, D. Temporal Change Detection in SAR Images Using Log Cumulants and Stacked Autoencoder. *IEEE Geosci. Remote Sens. Lett.* **2018**, *15*, 297–301. [CrossRef]

32. Zhang, X.; Su, X.; Yuan, Q.; Wang, Q. Spatial–Temporal Gray-Level Co-Occurrence Aware CNN for SAR Image Change Detection. *IEEE Geosci. Remote Sens. Lett.* **2022**, *19*, 4018605. [CrossRef]

33. Wang, R.; Wang, L.; Wei, X.; Chen, J.W.; Jiao, L. Dynamic Graph-Level Neural Network for SAR Image Change Detection. *IEEE Geosci. Remote Sens. Lett.* **2022**, *19*, 4501005. [CrossRef]

34. Mignotte, M. A Fractal Projection and Markovian Segmentation-Based Approach for Multimodal Change Detection. *IEEE Trans. Geosci. Remote Sens.* **2020**, *58*, 8046–8058. [CrossRef]

35. Sun, Y.; Lei, L.; Guan, D.; Wu, J.; Kuang, G.; Liu, L. Image Regression with Structure Cycle Consistency for Heterogeneous Change Detection. *IEEE Trans. Neural Netw. Learn. Syst.* **2022**, 1–15. [CrossRef]

36. Sun, Y.; Lei, L.; Tan, X.; Guan, D.; Wu, J.; Kuang, G. Structured graph based image regression for unsupervised multimodal change detection. *ISPRS J. Photogramm. Remote Sens.* **2022**, *185*, 16–31. [CrossRef]

37. Sun, Y.; Lei, L.; Guan, D.; Li, M.; Kuang, G. Sparse-Constrained Adaptive Structure Consistency-Based Unsupervised Image Regression for Heterogeneous Remote-Sensing Change Detection. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 4405814. [CrossRef]

38. Sun, Y.; Lei, L.; Guan, D.; Kuang, G. Iterative Robust Graph for Unsupervised Change Detection of Heterogeneous Remote Sensing Images. *IEEE Trans. Image Process.* **2021**, *30*, 6277–6291. [CrossRef]

39. Sun, Y.; Lei, L.; Li, X.; Sun, H.; Kuang, G. Nonlocal patch similarity based heterogeneous remote sensing change detection. *Pattern Recognit.* **2021**, *109*, 107598. [CrossRef]

40. Sun, Y.; Lei, L.; Guan, D.; Kuang, G.; Liu, L. Graph Signal Processing for Heterogeneous Change Detection. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 4415823. [CrossRef]

41. Zheng, Y.; Zhang, X.; Hou, B.; Liu, G. Using Combined Difference Image and k-Means Clustering for SAR Image Change Detection. *IEEE Geosci. Remote Sens. Lett.* **2014**, *11*, 691–695. [CrossRef]

42. Hou, B.; Wei, Q.; Zheng, Y.; Wang, S. Unsupervised Change Detection in SAR Image Based on Gauss–Log Ratio Image Fusion and Compressed Projection. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 3297–3317. [CrossRef]

43. Krähenbühl, P.; Koltun, V. Efficient inference in fully connected crfs with gaussian edge potentials. In Proceedings of the 21st Annual Conference on Neural Information Processing Systems, Granada, Spain, 12–17 December 2011.

44. 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]

45. Jimenez-Sierra, D.A.; Quintero-Olaya, D.A.; Alvear-Muñoz, J.C.; Benítez-Restrepo, H.D.; Florez-Ospina, J.F.; Chanussot, J. Graph Learning Based on Signal Smoothness Representation for Homogeneous and Heterogeneous Change Detection. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 4410416. [CrossRef]

46. Ban, Z.; Liu, J.; Cao, L. Superpixel Segmentation Using Gaussian Mixture Model. *IEEE Trans. Image Process.* **2018**, *27*, 4105–4117. [CrossRef] [PubMed]

47. Zheng, X.; Guan, D.; Li, B.; Chen, Z.; Li, X. Change Smoothness-Based Signal Decomposition Method for Multimodal Change Detection. *IEEE Geosci. Remote Sens. Lett.* **2022**, *19*, 2507605. [CrossRef]
48. Otsu, N. A threshold selection method from gray-level histograms. *IEEE Trans. Syst. Man Cybern.* **1979**, *9*, 62–66. [CrossRef]
49. Kittler, J.; Illingworth, J. Minimum error thresholding. *Pattern Recognit.* **1986**, *19*, 41–47. [CrossRef]
50. Hartigan, J.A.; Wong, M.A. Algorithm AS 136: A k-means clustering algorithm. *J. R. Stat. Soc. Ser. C (Appl. Stat.)* **1979**, *28*, 100–108. [CrossRef]
51. Bezdek, J.C.; Ehrlich, R.; Full, W. FCM: The fuzzy c-means clustering algorithm. *Comput. Geosci.* **1984**, *10*, 191–203. [CrossRef]
52. Sun, Y.; Lei, L.; Guan, D.; Wu, J.; Kuang, G. Iterative structure transformation and conditional random field based method for unsupervised multimodal change detection. *Pattern Recognit.* **2022**, *131*, 108845. [CrossRef]
53. Gong, M.; Cao, Y.; Wu, Q. A Neighborhood-Based Ratio Approach for Change Detection in SAR Images. *IEEE Geosci. Remote Sens. Lett.* **2012**, *9*, 307–311. [CrossRef]
54. Lei, L.; Sun, Y.; Kuang, G. Adaptive Local Structure Consistency-Based Heterogeneous Remote Sensing Change Detection. *IEEE Geosci. Remote Sens. Lett.* **2022**, *19*, 8003905. [CrossRef]