Signed Graph-Based Image Transformation for Heterogeneous Change Detection

Yuli Sun, Ming Li, Lin Lei, Zhang Li, and Gangyao Kuang, *Senior Member, IEEE*

Abstract— Heterogeneous change detection (HeCD) is a highly valuable yet challenging task in remote sensing. To enable the comparison of heterogeneous images with different imaging mechanisms, some structural consistency-based image transformation methods have been proposed, which utilize graph models to represent image structures and constrain the transformed images and original images to have the same structural characteristics on the graph model. Consequently, these graph-based methods face two challenges: adequately characterizing the image structure and effectively utilizing the change information. To address these challenges, this article proposes a signed graph-based image transformation (SGIT) method for unsupervised HeCD. First, we analyze the limitations of previous unsigned graph-based methods in capturing the image structure, which leads to the failure to detect changes in some scenes. In light of this, we construct signed graph models that utilize positive/negative weights to represent the similarity/dissimilarity relationships within the image, respectively, and employ adaptive weighting, negative sampling, and neighborhood expansion strategies to bolster the structure representation capability of signed graphs. Second, we analyze how the change would induce a bimodal distribution of vertex feature distances in original and transformed images. Subsequently, a distribution-induced reweighted graph Laplacian regularization (RGLR) is proposed to exploit this prior change information. Finally, a more accuracy image transformation model is obtained by incorporating three types of constraints: signed graph-based structural consistency term, bimodal distribution-induced RGLR, and change sparsity-based penalty term. Extensive comparative experiments on five real datasets have demonstrated the effectiveness of the proposed SGIT.

Index Terms— Change detection, heterogeneous data, image transformation, signed graph.

NOMENCLATURE

Symbol	Description
$\tilde{X}$, $\tilde{Y}$	Original pre- and post-even images.
$\tilde{X}'$, $\tilde{Y}'$	Transformed images.
$\tilde{X}_i$, $\tilde{Y}_i$	ith superpixels of original images.
X, Y	Feature matrices of original images.
$\mathbf{X}'$, $\mathbf{Y}'$	Transformed feature matrices.
Δ^x, Δ^y	Changed feature matrices.
G^{t1}, G^{t2}	Unsigned graphs.
$\mathcal{G}_\pm = (\mathcal{G}_+, \mathcal{G}_-)$	Signed graph with two subgraphs of $\mathcal{G}_+$ and $\mathcal{G}_-$.
$\mathcal{G}_{h\pm} = (\mathcal{G}_{h+}, \mathcal{G}_{h-})$	High-order signed graph.
$\mathbf{I}_N$	$N \times N$ identity matrix.

I. INTRODUCTION

A. Background

CHANGE detection refers to the technique of extracting change information by comparing multiple remote sensing images observed from the same region at different times [1]. Change detection has been widely used in fields such as environment monitoring, disaster evaluation, and agricultural surveys [2], [3].

Currently, most change detection methods utilize multiple images acquired under homogenous conditions with the same sensor, known as homogenous change detection (HoCD) or unimodal change detection. However, in certain practical scenarios, it is challenging or even impossible to obtain homogenous images before and after an event [4], [5], at this time, changes can only be detected with the help of heterogeneous images captured by different sensors, termed as heterogeneous change detection (HeCD) or multimodal change detection. For instance, in emergency response to some sudden disasters (earthquakes, floods, landslides, etc.), constrained by image resources and imaging conditions, analysts often rely on pre-event optical images and post-event synthetic aperture radar (SAR) images for change analysis [6], swiftly obtaining disaster information to support rescue efforts. In addition, in the analysis of changes over long time-series, by introducing heterogenous images in HeCD, the temporal range can be expanded and the temporal resolution of the analysis can be improved [7].

B. Related Work

Since the images used in HeCD come from different sensors (e.g., optical and SAR) with different imaging conditions, the heterogeneous images show great modal gap [8], [9] as illustrated in Fig. 1, making it impossible to identify changed region by directly comparing the pre- and post-event images as in the case of HoCD. Therefore, the first task of HeCD is to explore correlations between heterogeneous images and convert them into a common space to enable comparison.
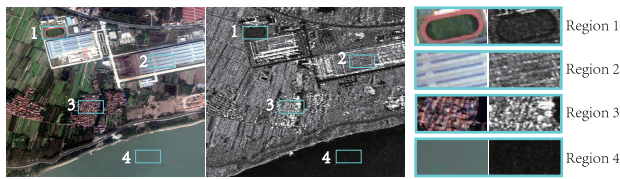
Fig. 1. Pair of optical-SAR images acquired from the same area. From these images and the zoomed-in subregions, it is clear that different sensors provide distinct perspectives on the same ground truth and exhibit different characteristics.

According to the comparative space, existing HeCD methods can be broadly categorized into three types: 1) image classification-based HeCD methods utilize classifiers to classify images into common land cover types and subsequently detect changes through comparing the categorization results, such as the hierarchical compound classification [10], and hierarchical extreme learning machine-based method [11]; 2) feature transformation-based HeCD methods assume some imaging-invariant correlations or utilize some deep Siamese/pseudo-Siamese networks to transform heterogeneous images into a shared feature space and then perform comparisons, such as the Copula theory-based similarity metrics [12], the spatial self-similarity difference-based method [13], self-guided autoencoders [14], hierarchical attention feature fusion-based network [15]; and 3) image transformation-based HeCD methods transform one image into another image domain, and then compare the original and transformed images within the same image domain, which is also referred to as style migration or image translation [16], [17], such as the fractal encoding–decoding [18], traditional image regression based on affinity matrix difference (AMD) [19], copula mixtures-based networks [20], code-aligned autoencoders [21], AMD-guided deep translation network [22], and multidomain constrained translation network [23].

Existing HeCD methods either rely on certain imaging-invariant assumptions or use labeled samples to train networks to establish correlations between heterogeneous images. Therefore, the robustness of the established correlations is very important for these methods. Recently, similarity graph-based HeCD methods have attracted the attention of researchers, which utilize the structural consistency between heterogeneous images to transform them into the same feature domain or image domain [24]. For example, some structure comparison-based HeCD methods construct or learn structural graphs to characterize the similarity relationships within images, and then compare the structural differences by graph mapping, such as the nonlocal patch graph comparison method [4], local and nonlocal structure graph fusion-based method [25], iterative robust graph-based method [26], adaptive optimization of a structured graph [27], and structural graph representation learning-based method [28], enhanced graph structure representation-based method [29]. Some structure regression-based HeCD methods utilize structural graphs to transform one image to the other image domain using the structural consistency constraint between original and transformed images, such as the image translation methods based on structure regression [30],

change smoothness [31], and similarity graph-based image transformation network [32].

These structural graph-based HeCD methods present three advantages: 1) these methods are intuitive and interpretable; 2) since the image structure exhibits the property of imaging-invariant, these methods are ideal for the unsupervised HeCD problem as they can establish the correlation between heterogeneous images without labeling samples; and 3) the image structure is less disturbed by noise than image pixel value, which enhances the robustness of these methods. While existing structural graph-based HeCD methods perform relatively well in HeCD, they still suffer from the following two shortcomings.

First, the graph model could not adequately characterize the image structural information, which reduces the performance of the structure comparison and structure transformation algorithms, resulting in limited change detection accuracy. Previous methods all focus solely on the use of unsigned graph models that characterize the similarity relationships of the image, while neglecting to explore and utilize the dissimilarity relationships within the image, such as the k-nearest neighbors (KNNs) graphs used in the INLPG [4], IRG-McS [26], AOSG [27], the high-order similarity graph in [25], the KNN graph-based hypergraph in [30], the adaptive local structure graph in [33], the graph convolutional autoencoder-based local and nonlocal similarity graphs [28], and the graph convolutional network in [32]. For structural graph-based methods, the representational capacity of the graph is a crucial factor affecting change detection performance. If the graph model can more fully extract image structure information, then the constraints based on structural consistency used in the structural comparison and structural regression models will be more precise, thereby improving the HeCD performance.

Second, previous methods failed to fully exploit the change prior information inherent in the HeCD problem, consequently limiting the expressive capability of the image transformation model. This limitation makes it challenging to effectively utilize prior knowledge and accurately extract change information. Currently, most methods leverage the change information in the HeCD problem primarily in two key aspects: spatial smoothness and sparsity of change. For instance, some methods consider the change smoothness, that is, spatially neighboring pixels tend to belong to the same class of objects, and therefore their change states tend to be the same as well. This concept is often employed during the segmentation of difference image (DI) to derive binary change map (CM) [26], [28], [31]. Concurrently, some methods incorporate change sparsity as a constraint in the image transformation model, that is, the proportion of regions experiencing change is minimal, with the majority remaining unchanged throughout the event [32], [33]. However, in addition to the spatial smoothness and sparsity, there are other prior knowledge about changes that can be mined in the HeCD problem. Further utilization of this knowledge can lead to better change detection results.

C. Motivations and Contributions

In this article, we first analyze the shortcomings of previous unsigned graph models in image structure-based HeCD

methods. Subsequently, we emphasize the advantages of the signed graph model in characterizing image structure and HeCD methods. With this perspective, we develop a signed graph model for each image and propose a signed graph-based image transformation (SGIT) method for HeCD. Specifically, we employ signed weights to depict the similarity/dissimilarity relationships within an image. That is, positive weights connect similar superpixels (representing the same class of objects), while negative weights link dissimilar superpixels (representing different classes of objects), reflecting attractive and repulsive forces between superpixels within an image, respectively. Subsequently, the signed graphs are applied to the image transformation model to constrain the transformed images and original images to exhibit the same similarity/dissimilarity relationships, i.e., possess identical structural properties on the signed graphs, which yields more precise transformed images and DIs than using only the unsigned graphs.

Furthermore, we delve deeper into the change prior to HeCD problem and identify differences in the distribution of vertex feature distances between pre- and post-event images as a consequence of changes. Building on this observation, we propose a reweighted graph Laplacian regularization (RGLR) in the image transformation model. Specifically, we utilize an image decomposition model to decompose the target domain image into a transformed image derived from the source domain and a change image caused by the change event, and employ three constraints: 1) a signed graph-based structural consistency term that constrains the transformed image and source domain image to have the same similarity/dissimilarity relationships on the signed graph; 2) a change sparsity term that constrains the change image to be sparse; and 3) a distribution-induced reweighted graph Laplacian term that can efficiently promote a bimodal distribution of vertex feature distances for the transformed image. By incorporating these three constraints, a more accurate and robust image transformation model can be obtained than the previous methods that utilized only the unsigned graph constraint and the change sparsity constraint, leading to better performance in the HeCD. The main contributions of this article are as follows.

1) We proposed a signed graph model-based image transformation method for the HeCD problem. This approach involves constructing signed graphs to represent the similarity/dissimilarity relationships among original images and subsequently ensuring that the transformed and original images exhibit identical structural properties on the signed graph. To the best of our knowledge, this is the first method to introduce signed graphs into a change detection task, addressing the limitations of previous unsigned graphs-based HeCD methods.

2) We conducted an analysis to elucidate how changes can induce a bimodal distribution of vertex feature distances in the original and transformed images. Subsequently, we designed a distribution-induced RGLR for the image transformation model, which enhances the accuracy and robustness of the model.

3) We introduced a robust signed graph construction method by leveraging adaptive weighting, negative sam-

pling, and neighborhood expansion strategies, which bolster the ability of the signed graph to characterize the similarity/dissimilarity structure information of the image effectively.

II. LIMITATIONS OF UNSIGNED GRAPH IN THE HeCD

A. Problem Formulation and Structural Consistency

Given two co-registered images collected by different sensors at times t_1 (pre-event) and t_2 (post-event), denoted as $\tilde{X} \in \mathbb{R}^{M \times N \times B_1}$ and $\tilde{Y} \in \mathbb{R}^{M \times N \times B_2}$, where the pixels are defined as $\tilde{x}_{m,n,b_1}$ and $\tilde{y}_{m,n,b_2}$, respectively. We define the feature matrix of the original pre-event image as $\mathbf{X}$, whose element $x_{i,j}$ represents the ith feature of the jth pixel (or patch or superpixel depending on the underlying processing unit). Similarly, we can also define the feature matrix $\mathbf{Y}$ for the post-event image. The HeCD task is to enable the comparison of heterogeneous images $\tilde{X}$ and $\tilde{Y}$ to compute a CM indicating the changes. The nomenclature section lists some important notations for easy reference.

Image transformation-based HeCD methods first transform one image to the domain of the other image, defined as $\mathcal{M}_1 : \tilde{X} \rightarrow \tilde{Y}'$ and $\mathcal{M}_2 : \tilde{Y} \rightarrow \tilde{X}'$, where $\tilde{X}'$ and $\tilde{Y}'$ are the transformed images, i.e., the assumed unchanged images. Then, the changed regions can be detected by comparing the original and transformed images, such as $\tilde{X} - \tilde{X}'$ and $\tilde{Y} - \tilde{Y}'$.

To obtain the transformed images, some constraints are added to the image transformation process, such as structural consistency constraints. Structural consistency between heterogeneous images is referred to as follows: although the heterogeneous images acquired by different sensors in the same area are distinct in terms of their presentation, the topological relationships between objects inside the images are consistent [4]. Structural consistency is rooted in the imaging mechanism of remote sensing, signifying that the same object maintains the same topological relationships across images captured under different conditions as long as the inherent properties of the objects remain unchanged. For example, suppose that the multitemporal images are divided into a series of superpixels (or square patches), denoted as $\tilde{X}_i$, $\tilde{Y}_i$, $i = 1, \ldots, N_S$, then if both the i and jth superpixels are unchanged during the event, their similarity/dissimilarity relationships can be well preserved across the heterogeneous images. That is, when $\tilde{X}_i$ and $\tilde{X}_j$ remain unchanged, and if they represent the same or different kinds of objects in the pre-event image, showing that $\tilde{X}_i$ and $\tilde{X}_j$ are very similar or dissimilar, then $\tilde{Y}_i$ and $\tilde{Y}_j$ also represent the same or different kinds of objects in the post-event image, showing that $\tilde{Y}_i$ and $\tilde{Y}_j$ are also very similar or not.

B. Unsigned Graph-Based Image Transformation Method

Using the structural consistency property, some image transformation-based HeCD methods have been proposed [31], [32], [33]. In these methods, they first construct unsigned graph (e.g., KNN graphs) of G^{t1} (or G^{t2}) for pre-event image (or post-event image) that connect each superpixel $\tilde{X}_i$ (or $\tilde{Y}_i$) with its similar neighbors $\tilde{X}_j$ (or $\tilde{Y}_j$) with the positive weight W_{ij}^{t1} (or W_{ij}^{t2}). Then, they decompose the original
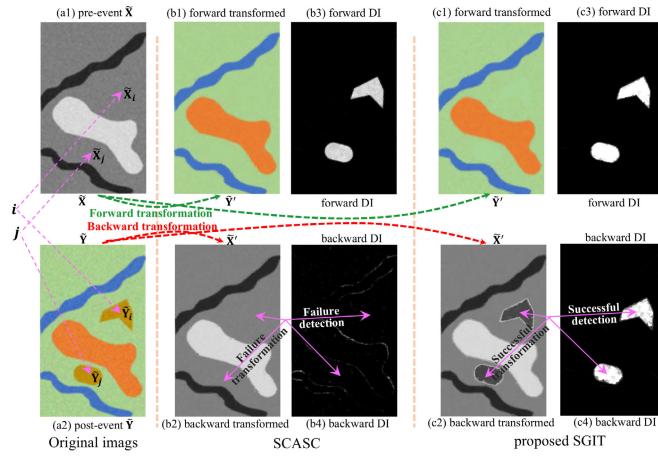
Fig. 2. HeCD example of SCASC [33] and proposed SGIT. (a1) and (a2) Original pre- and post-event images. (b1)–(b4) Detection results generated by SCASC: forward transformed image, backward transformed image, forward DI, and backward DI, respectively. (c1)–(c4) Detection results generated by the proposed SGIT. During the forward transformation process, both SCASC and SGIT are able to transform images and detect changes. However, in the backward transformation process, only SGIT is capable of detecting changes, while SCASC cannot, because SGIT uses a signed graph instead of an unsigned graph.

feature matrices ($\mathbf{Y}$ and $\mathbf{X}$) into transformed feature matrices ($\mathbf{Y}'$ and $\mathbf{X}'$) and changed feature matrices (Δ^y and Δ^x), denoted as $\mathbf{Y} = \mathbf{Y}' - \Delta^y$ and $\mathbf{X} = \mathbf{X}' - \Delta^x$, and named as forward transformation (1) and backward transformation (2), respectively,

$$\min_{\mathbf{Y}', \Delta^y} \sum_{i,j=1}^{N_S} W_{ij}^{t1} \left\| \mathbf{Y}_i' - \mathbf{Y}_j' \right\|_2^2 + \lambda \left\| \Delta^y \right\|_{2,1} \text{ s.t. } \mathbf{Y}' = \mathbf{Y} + \Delta^y \quad (1)$$

$$\min_{\mathbf{X}', \Delta^x} \sum_{i,j=1}^{N_S} W_{ij}^{t2} \left\| \mathbf{X}_i' - \mathbf{X}_j' \right\|_2^2 + \lambda \left\| \Delta^x \right\|_{2,1} \text{ s.t. } \mathbf{X}' = \mathbf{X} + \Delta^x \quad (2)$$

where $\lambda > 0$ is the balance parameter, and the $\ell_{2,1}$-norm of $\|\Delta^y\|_{2,1} = \sum_{i=1}^{N_S} \|\Delta_i^y\|_2$ and $\|\Delta^x\|_{2,1} = \sum_{i=1}^{N_S} \|\Delta_i^x\|_2$ represent the change sparsity regularization.

In the transformation models of (1) and (2), the first terms of $\sum_{i,j=1}^{N_S} W_{ij}^{t1} \|\mathbf{Y}_i' - \mathbf{Y}_j'\|_2^2$ and $\sum_{i,j=1}^{N_S} W_{ij}^{t2} \|\mathbf{X}_i' - \mathbf{X}_j'\|_2^2$ are used to constrain the transformed images ($\tilde{\mathbf{Y}}'$, $\tilde{\mathbf{X}}'$) and original images ($\tilde{\mathbf{X}}$, $\tilde{\mathbf{Y}}$) to have the same similarity structure on the unsigned graphs (G^{t1}, G^{t2}). The second terms of $\|\Delta^y\|_{2,1}$ and $\|\Delta^x\|_{2,1}$ are used to constrain the changed image to be sparse in piratical HeCD problem. After completing the transformation processes, the Δ^y and Δ^x can be segmented to extract the changed regions.

C. Limitations of Unsigned Graph-Based Transformation Method

Fig. 2 gives a simple example of unsigned graph-based SCASC [33] and the proposed signed graph base SGIT, from which it can be seen that SCASC effectively generates transformed images and DIs in the forward transformation process. However, it encounters difficulties in the backward transformation. Importantly, this phenomenon is not unique to SCASC, there are other methods that also suffer from

such issues of succeeding in one transformation process but failing in the other, such as AMDIR [19] and ACE-Net [22], as evidenced by our experimental results in the Section IV.

This diametrically different performance of forward and backward transformations has also been noticed in recent literature [30], [32]. Florez-Ospina et al. [32] found a negative impact on change detection results when reversing the input order of pre- and post-event images, prompting them to inquire in the concluding section: "*why is this happening in the first place, and how can we maintain performance while making the method invariant to such reversal?*" In our previous study [30], we have attributed this reversal inconsistency to the structural asymmetry between heterogeneous images, i.e., the structural complexities of heterogeneous images are different, resulting in the failure of transforming from a complex structural image to a simpler one. Furthermore, the recent work [34] has analyzed this phenomenon from the perspectives of the transformation model itself and the frequency property of the KNN graph. Contrary to previous studies, this article develops an analysis of the difference between unsigned and signed graphs under the perspective of graph signal processing (GSP), and then proposes a signed graph-induced transformation model, which is more concise and generalizable. We believe that our findings could well address the query posted by Florez-Ospina et al. [32].

We take the failed backward transformation in Fig. 2 as an example for the limitation analysis. In the unsigned graph G^{t2}, it uses the positive weight W_{ij}^{t2} to describe the "similarity" or "attraction" relationships between the internal superpixels within the image $\tilde{\mathbf{Y}}$. For example, for the KNN graph, the vertices connected by its edges represent the same class of objects (when an appropriate k is chosen); for the fully connected unsigned graph, its edge weight W_{ij}^{t2} reflect the similarity between vertices of $\tilde{\mathbf{Y}}_i$ and $\tilde{\mathbf{Y}}_j$, typically with $W_{ij}^{t2} = \exp(-\|\mathbf{Y}_i - \mathbf{Y}_j\|_2^2/\sigma)$. Based on the structural consistency property, it requires that the transformed image (i.e., the unchanged image) $\tilde{\mathbf{X}}'$ share the same relationship as the original post-event image $\tilde{\mathbf{Y}}$. Therefore, in the unsigned graph-based backward transformation model (2), it uses the regularization term of $\sum_{i,j=1}^{N_S} W_{ij}^{t2} \|\mathbf{X}_i' - \mathbf{X}_j'\|_2^2$ to constrain that the connected vertices of $\tilde{\mathbf{X}}_i'$ and $\tilde{\mathbf{X}}_j'$ by graph G^{t2} have the same similarity relationship as $\tilde{\mathbf{Y}}_i$ and $\tilde{\mathbf{Y}}_j$.

If we treat the transformed image as a graph signal on the defined graph, whose feature vector $\mathbf{X}_i$ denotes the signal on the ith vertex, then the regularization term of $\sum_{i,j=1}^{N_S} W_{ij}^{t2} \|\mathbf{X}_i' - \mathbf{X}_j'\|_2^2$ is equal to the total variation (TV) on the graph G^{t2} [35], which can be regarded as the smoothness penalty term used to constrain the smoothness of the transformed image $\tilde{\mathbf{X}}'$ on the graph G^{t2}. However, when the pre-event image $\tilde{\mathbf{X}}$ itself is a smooth signal on G^{t2}, and taking into account the sparse constraint term, then the output of backward transformation model (2) at this point is $\Delta^x = \mathbf{0}$ and $\mathbf{X}' = \mathbf{X}$. As shown in Fig. 2, for any $\tilde{\mathbf{Y}}_i$ and $\tilde{\mathbf{Y}}_j$ in the post-event image that represents the same class of objects, their corresponding $\tilde{\mathbf{X}}_i$ and $\tilde{\mathbf{X}}_j$ in the pre-event image also represent the same class of objects, implying that $\tilde{\mathbf{X}}$ is smooth on graph G^{t2}, which thus leads to the failure of

SCASC on backward transformation. Furthermore, consider an extreme case when the pre-event images entirely represent the same class of objects (completely smooth on any unsigned graph), i.e., $\mathbf{X}_i = \mathbf{X}_j$, $\forall i, j \in \{1, \cdots N_S\}$, then the output of model (2) is always $\Delta^x = \mathbf{0}$ and $\mathbf{X}' = \mathbf{X}$, which means that it cannot detect any change at this point, no matter what the post-event image is and what changes have occurred. It can be concluded that when the unsigned graph model (e.g., G^{t2}) constructed on one image (e.g., $\tilde{Y}$) fails to distinguish between two heterogeneous images, the transformation model fails in that direction (e.g., $\mathcal{M}_2 : \tilde{Y} \to \tilde{X}'$).

From the above analysis, it can be seen that the reason for the transformation failure is that only the unsigned graph-induced smoothness penalty of the TV regularization term (i.e., smoothness on the graph) is utilized, which only exploits the similarity relationships in the structural consistency between the heterogeneous images, but neglect the dissimilarity consistency (i.e., variability on the graph). Opposite to the "similarity" or "attraction" relationships within the image is the notion of "dissimilarity" or "repulsion." As shown in Fig. 2, although $\tilde{X}$ exhibits the same similarity relationships as $\tilde{Y}$ in the graph G^{t2}, its dissimilarity relationships differ significantly from $\tilde{Y}$, which can be leveraged to detect the changes.

In this article, when two superpixels represent different kinds of objects, showing that they are very dissimilar, then we model this "dissimilarity" with a negative edge. The meaning of a negative edge is NOT any edge or edge with a small positive weight; rather, the negative edge implies the connected samples are expected to have different values. Intuitively, the penalty term $\sum_{i,j=1}^{N_S} W_{ij}^{t2} \|\mathbf{X}'_i - \mathbf{X}'_j\|_2^2$ with zero (means no edge) or very small positive weight W_{ij}^{t2} has little constraint on the connected $\mathbf{X}'_i$ and $\mathbf{X}'_j$. Conversely, the penalty term $\sum_{i,j=1}^{N_S} W_{ij}^{t2} \|\mathbf{X}'_i - \mathbf{X}'_j\|_2^2$ with negative weight W_{ij}^{t2} will compel $\mathbf{X}'_i$ and $\mathbf{X}'_j$ to diverge from each other. By incorporating the pairwise dissimilarity, the signed graph has achieved enhanced performance across various fields, such as data mining [36] and social relations analyzing [37]. In this article, we leverage a signed graph to address the transformation failure or degradation issue due to the unsigned graph, as shown in Fig. 3.

III. SIGNED GRAPH-BASED IMAGE TRANSFORMATION METHOD FOR HeCD

A. Preprocessing

Since pairwise similarity/dissimilarity relationships need to be considered, we opt for superpixel as the basic unit in the proposed SGIT. This choice offers two main advantages: first, it significantly reduces the number of graph vertices, thereby decreasing the algorithm complexity; second, it can effectively utilize the context information and retain the edge contours of objects.

To segment the multitemporal images into superpixels, we perform the superpixel co-segmentation procedure as in [30]. It first constructs a pseudo-RGB image, where the first two channels are the normalized gray values of the multitemporal images, and the third channel is all zero-valued. Then, it employs the GMMSP [38] to segment the
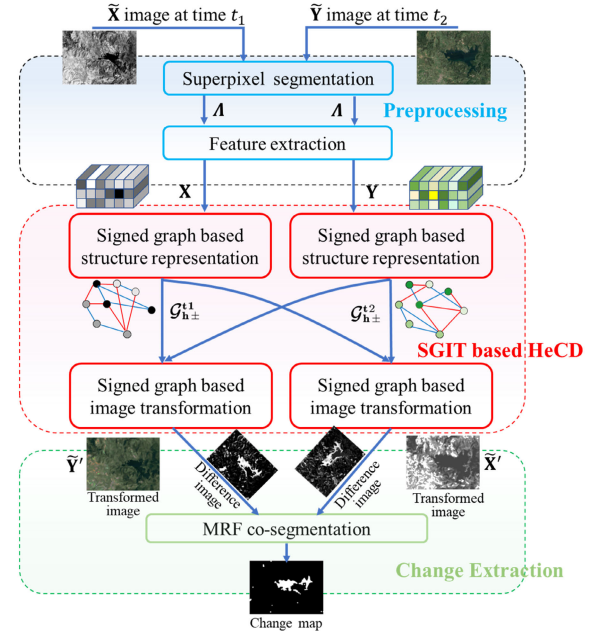

Fig. 3. Framework of the SGIT for HeCD.

pseudo-RGB image into superpixels with the segmentation map Λ. By mapping Λ into the multitemporal images, the co-segmented superpixels of $\tilde{X}_i$ and $\tilde{Y}_i$, $i = 1, 2, \ldots, N_S$ are generated as

$$\tilde{X}_i = \left\{ \tilde{x}_{m,n,b_1} | (m,n) \in \Lambda_i, b_1 = 1, \ldots, B_1 \right\}$$
$$\tilde{Y}_i = \left\{ \tilde{y}_{m,n,b_2} | (m,n) \in \Lambda_i, b_2 = 1, \ldots, B_2 \right\}. \quad (3)$$

In this manner, $\tilde{X}_i$ and $\tilde{Y}_i$ belong to the same geographic area and are each homogeneous inside with high probability. After the superpixel co-segmentation, the median and mean values are extracted as superpixel features to compose the feature matrices of $\mathbf{X} \in \mathbb{R}^{2B_1 \times N_S}$ and $\mathbf{Y} \in \mathbb{R}^{2B_2 \times N_S}$.

B. Structure Representation by Signed Graph

Previous structure-based HeCD methods focus on the use of unsigned graph models to characterize the similarity relationships of the image, such as the KNN graph [27], self-similarity graph [33], adaptive local structure graph [33], and the graph convolutional network [32], [39]. Here, we utilize a signed graph to capture image structure, which uses positive edge weights to portray similarity relationships and negative edge weights to portray dissimilarity relationships within the image, respectively.

We define the signed graph as $\mathcal{G}_\pm = (\mathcal{G}_+, \mathcal{G}_-)$, where $\mathcal{G}_+ = (\mathcal{V}, \mathcal{E}_+)$ and $\mathcal{G}_- = (\mathcal{V}, \mathcal{E}_-)$ are two subgraphs that encode positive and negative links, respectively, with a common vertex set $\mathcal{V} = \{1, 2, \ldots, N_S\}$. The weighting matrix $\mathbf{W}_\pm$ of the signed graph can be decomposed into $\mathbf{W}_\pm = \mathbf{W}_+ + \mathbf{W}_-$ with

$$W_{+,ij} = \begin{cases} W_{\pm,ij}, & \text{if } W_{\pm,ij} > 0 \\ 0, & \text{otherwise} \end{cases}$$

and

$$W_{-,ij} = \begin{cases} W_{\pm,ij}, & \text{if } W_{\pm,ij} < 0 \\ 0, & \text{otherwise.} \end{cases}$$

1) Graph $\mathcal{G}_+$ With Positive Edge Weights: In the signed graph $\mathcal{G}_\pm$, it uses the positive edge weight to describe the "similarity" or "attraction" relationships between superpixels, i.e., the subgraph of $\mathcal{G}_+$. Starting with the pre-event image $\tilde{X}$ collected at time t_1, we initially build a probabilistic graph $\mathcal{G}_+^{t1}$ that connects each superpixel with similar neighbors with positive weight matrix $\mathbf{W}_+^{t1}$, which follows the model used in SCASC [33] as:

$$\min_{\mathbf{W}_+^{t1}} \sum_{i,j=1}^{N_S} \text{dist}_{ij}^x W_{+,ij}^{t1} + \alpha \left(W_{+,ij}^{t1}\right)^2$$

$$\text{s.t.} \sum_{j=1}^{N_S} W_{+,ij}^{t1} = 1, \quad 0 \leq W_{+,ij}^{t1} \leq 1 \tag{4}$$

where dist_{ij}^x represents the feature distance with $\text{dist}_{ij}^x = \|\mathbf{X}_i - \mathbf{X}_j\|_2^2$, and $\alpha > 0$ is a control parameter that determines the number of neighbors. The closed-form of $\mathbf{W}_+^{t1}$ can be obtained using the solution method in [33] with

$$W_{+,i(j)}^{t1} = \begin{cases} \dfrac{\text{dist}_{i(k+1)}^x - \text{dist}_{i(j)}^x}{k\text{dist}_{i(k+1)}^x - \sum_{h=1}^{k} \text{dist}_{i(h)}^x}, & j \leq k \\ 0, & j > k \end{cases} \tag{5}$$

where $\text{dist}_{i(j)}^x$ denotes the jth smallest distance in the vector $\text{dist}_i^x = \{\text{dist}_{ij}^x | j = 1, \ldots, N_S\}$, and the symbol (j) in $W_{+,i(j)}^{t1}$ and $\text{dist}_{i(j)}^x$ denotes the index of the jth smallest value.

Therefore, it can be noted that $\mathcal{G}_+^{t1}$ is actually also a KNN graph model. To choose an appropriate number of neighbors k for each superpixel, we use the following steps similar to [33].

1) Define the maximum and minimum number of possible neighbors as $k_{\max} = \lceil \sqrt{N_S} \rceil$ and $k_{\min} = \lceil \sqrt{N_S}/10 \rceil$, where $\lceil \cdot \rceil$ denotes the upward rounding. Then, we find the $k_{\max}$ nearest neighbors for each superpixel.
2) 2) Compute the in-degree $\text{Id}(i)$ for the ith superpixel, i.e., the number of times of $\tilde{X}_i$ appears in the $k_{\max}$-nearest neighbors of other superpixels.
3) Finally, the number of neighbors for each superpixel is set to be $k_i = \min\{k_{\max}, \max\{k_{\min}, \text{Id}(i)\}\}$. This data-dependent neighbors selection and edge weights calculation (5) can help the subgraph $\mathcal{G}_{t1}^+$ adaptively capture similarity relationships. Similarly, we can construct the probabilistic graph $\mathcal{G}_{t2}^+$ for the post-event image $\tilde{Y}$ with the same step as $\mathcal{G}_{t1}^+$, and obtain the positive edge weights $\mathbf{W}_+^{t2}$.

2) Graph $\mathcal{G}_-$ With Negative Edge Weights: In the signed graph, it uses the negative edge weight to describe the "dissimilarity" or "repulsion" relationships between superpixels, i.e., the subgraph of $\mathcal{G}_-$. In the recent previous work of [34], we have proposed a k-farthest neighbors (KFNs) graph to capture the dissimilarity relationships within the image. For example, in the pre-event image, the KFN graph $G^{t1,f}$ connects each superpixel with its k-most different superpixels with weight $W_{i(j)}^{t1,f}$. That is, if and only if distance $\text{dist}_{i(j)}^x$ belongs to the k-largest values in the vector dist_i^{t1}, then the jth superpixel will be connected with the ith superpixel, i.e.,

$$W_{i(j)}^{t1,f} = \begin{cases} 1, & N_S - k < j \leq N_S \\ 0, & j \leq N_S - k. \end{cases}$$

By selecting the k-most different superpixels as neighbors, the KFN graph is able to guarantee that the vertices connected by its edges are all representing different kinds of objects. However, this KFN graph also introduces a very serious drawback: each vertex is likely to be connected to only a very small number of classes or even to the same class of superpixels. In other words, the ith superpixel connected k-most different superpixels of $\{\tilde{X}_j | W_{i(j)}^{t1,f} = 1, j \in \mathcal{V}\}$ have a high probability of belonging to the same kind of objects. This limitation hampers the graph model's capacity to delineate dissimilarity relationships effectively, as our primary intention is to have each vertex connected to vertices representing a diversity of different kinds of objects.

Different from the previous KFN graph, we employ the negative sampling strategy to construct the dissimilarity graph $\mathcal{G}_-^{t1}$, which consists of two steps: 1) determine the range of candidate dissimilar neighbors for the ith superpixel as $\mathcal{C}_i = \{\tilde{X}_{(j)} | q < j \leq N_S\}$, with (j) denoting the index of the jth smallest value in distance vector dist_i^x and q being empirically set to $2N_S/3$ and 2) set the sampling probability of candidate neighbor superpixels for the ith superpixel as

$$\begin{aligned} &p_{i(j)} \\ &= \begin{cases} \dfrac{\text{dist}_{i(j)}^x - \text{dist}_{i(q)}^x}{\sum_{n=q+1}^{N_S} \text{dist}_{i(n)}^x - (N_S - q)\text{dist}_{i(q)}^x}, & N_S - q < j \leq N_S \\ 0, & j \leq N_S - q. \end{cases} \end{aligned} \tag{6}$$

Then, we randomly sample $k_{\max}$ superpixels from $\mathcal{C}_i$ with this $p_{i(j)}$ as the neighbor to construct negative edge with weight $W_{-,i(j)}^{t1} = -1/k_{\max}$. This negative sampling-based neighbor selection strategy can help each vertex connect vertices representing different classes of objects, which improves the graph model's ability to portray dissimilarity relationships. In a similar manner, we can build the dissimilarity graph $\mathcal{G}_-^{t2}$ for the post-event image, and obtain the negative edge weights $\mathbf{W}_-^{t2}$.

3) High-Order Signed Graph: To capture the high-order similarity/dissimilarity structure feature of image, we build high-order signed graphs $\mathcal{G}_{h\pm} = (\mathcal{G}_{h+}, \mathcal{G}_{h-})$ based on the above subgraphs of $\mathcal{G}_+$ and $\mathcal{G}_-$. We propose two neighborhood expansion strategies based on the well-known balance theory in social psychology [40], [41], which posits that social relationships adhere to four rules: "the friend of my friend is my friend," "the enemy of my friend is my enemy," "the friend of my enemy is my enemy," and "the enemy of my enemy is my friend."

For the high-order subgraph $\mathcal{G}_{h+}$, we use the neighborhood expansion strategy of "the neighbors of neighbors are also considered as neighbors" similar to the first rule in balance theory, which helps to construct a high-order probabilistic graph connecting more pairwise similar superpixels truly representing the same kind of objects. For example, the second-order subgraph $\mathcal{G}_{2+}^{t1} = (\mathcal{V}, \mathcal{E}_{2+}^{t1})$ is defined with

$$\begin{aligned} (i, j) \in \mathcal{E}_{2+}^{t1}; \quad &\text{if} \quad (i, j) \in \mathcal{E}_+^{t1} \\ &\text{or} \quad (i, u) \in \mathcal{E}_+^{t1}, \quad (u, j) \in \mathcal{E}_+^{t1}. \end{aligned} \tag{7}$$

The weighting matrix of h-order graph $\mathcal{G}_{h+}^{t1}$ is computed as

$$\mathbf{W}_{h+}^{t1} = \left(\mathfrak{D}_{h+}^{t1}\right)^{-1} \odot \mathfrak{W}_{h+}^{t1} \tag{8}$$

where $\mathfrak{W}_{h+}^{t1} = \sum_{n=1}^{h}(\mathbf{W}_{+}^{t1})^n$ and $\mathfrak{D}_{h+}^{t1}$ is a diagonal matrix with diagonal elements of $\mathfrak{D}_{h+,ii}^{t1} = \sum_{j=1}^{N_S} \mathfrak{W}_{h+,ij}^{t1}$.

For the high-order subgraph $\mathcal{G}_{h-}$, we use the neighborhood expansion strategies of "*the dissimilar-neighbors of the similar-neighbors are also dissimilar-neighbors*" and "*the similar-neighbors of the dissimilar-neighbors are also dissimilar-neighbors*" like the second and third rules in balance theory, respectively, which helps each vertex to connect more dissimilar vertices that represent different kinds of objects. For example, the second-order subgraph of $\mathcal{G}_{2-}^{t1} = (\mathcal{V}, \mathcal{E}_{2-}^{t1})$ is defined with

$$
\begin{aligned}
(i, j) \in \mathcal{E}_{2-}^{t1}; \quad &\text{if} \quad (i, j) \in \mathcal{E}_{-}^{t1} \\
&\text{or} \quad (i, u) \in \mathcal{E}_{-}^{t1}, \ (u, j) \in \mathcal{E}_{+}^{t1} \\
&\text{or} \quad (i, u) \in \mathcal{E}_{+}^{t1}, \ (u, j) \in \mathcal{E}_{-}^{t1}.
\end{aligned} \tag{9}
$$

The weighing matrix of h-order graph $\mathcal{G}_{h-}^{t1}$ is computed as

$$\mathbf{W}_{h-}^{t1} = -\left(\mathfrak{D}_{h-}^{t1}\right)^{-1} \odot \mathcal{A}_{h-}^{t1} \tag{10}$$

where $\mathcal{A}_{h-}^{t1}$ denotes the adjacent matrix, $\mathfrak{D}_{h-}^{t1}$ denotes the diagonal degree matrix, and $\mathbf{W}_{h-}^{t1}$ represents the random walk matrix of the graph $\mathcal{G}_{h-}^{t1}$.

Using these two neighborhood expansion strategies, we obtain the high-order signed graph of $\mathcal{G}_{h\pm}^{t1} = (\mathcal{G}_{h+}^{t1}, \mathcal{G}_{h-}^{t1})$, which can connect more similar and dissimilar neighbors with positive and negative weights, respectively. In a similar manner, we build the high-order signed graph $\mathcal{G}_{h\pm}^{t2} = (\mathcal{G}_{h+}^{t2}, \mathcal{G}_{h-}^{t2})$ for the post-event image, and obtain the edge weights of $\mathbf{W}_{h+}^{t2}$ and $\mathbf{W}_{h-}^{t2}$, whose the positive and negative edge weights are exploited to capture the attraction and repulsion relationships within the image, respectively.

Furthermore, it is noteworthy that we do not use the fourth rule of balance theory in our neighborhood expansion strategies. This is due to the fact that in remote sensing images, there are numerous object classes (much more than just two), so we can't judge whether the dissimilar-neighbor of the vertex's dissimilar-neighbor is a similar-neighbor or a dissimilar-neighbor (i.e., whether it belongs to the same object class as the vertex).

C. Signed Graph-Based Transformation Model

For image transformation-based HeCD methods, they transform image from one domain to the other domain to make heterogeneous image comparable, defined as $\mathcal{M}_1 : \tilde{X} \rightarrow \tilde{Y}'$ and $\mathcal{M}_2 : \tilde{Y} \rightarrow \tilde{X}'$. In the proposed SGIT, since the superpixel features are employed for representing the images, it is necessary to establish the transformation function between feature matrices as $\mathcal{T}_1 : X \rightarrow Y'$ and $\mathcal{T}_2 : Y \rightarrow X'$, with Y' and X' denoting the transformed feature matrices. By defining the operator for extracting the superpixel feature as $\mathcal{F}$ and the operator for recovering pixel values from the feature as $\mathcal{F}^{-1}$, the image transformation process can be described as

$$\text{Forward} \quad \mathcal{M}_1 : \tilde{Y}' = \mathcal{F}^{-1} \mathcal{T}_1 \mathcal{F}(\tilde{X})$$

$$\text{Backward} \quad \mathcal{M}_2 : \tilde{X}' = \mathcal{F}^{-1} \mathcal{T}_2 \mathcal{F}(\tilde{Y}). \tag{11}$$

Next, we detail these two transformations.

1) Signed Graph-Based SCR: Since the transformed images can be regarded as the assumed unchanged images, they retain the similarity/dissimilarity relationships within the original images based on the structural consistency property between the heterogeneous images. Therefore, the transformed images $\tilde{X}'$ and $\tilde{Y}'$ have the same property with the original images $\tilde{Y}$ and $\tilde{X}$ on the signed graph $\mathcal{G}_{h\pm}^{t2}$ and $\mathcal{G}_{h\pm}^{t1}$, respectively.

Take the forward transformation (11) as an example, according to the signed graph construction process of $\mathcal{G}_{h\pm}^{t1}$, if the superpixel pairs $\tilde{X}_i$ and $\tilde{X}_j$ in the original image are connected with positive edge $W_{h+,ij}^{t1}$ (i.e., connected by subgraph $\mathcal{G}_{h+}^{t1}$), then they are most likely to belong to the same kind of objects, and accordingly the corresponding superpixel pairs $\tilde{Y}_i'$ and $\tilde{Y}_j'$ in the transformed image also belong to the same kind of objects, and the difference between their feature vectors of $Y_i - Y_j$ should be very small. On the contrary, if the superpixel pairs in the original image are connected with the negative edge (i.e., connected by subgraph $\mathcal{G}_{h-}^{t1}$), then they are most likely to belong to the different kinds of objects, and accordingly the corresponding superpixel pairs in the transformed image also belong to the different kinds of objects, and showing very different feature vectors.

In Appendix A, we analyze that previous conventional signed graph smoothness constraints cannot be used directly in the image transformation-based HeCD. In this article, we use the following signed graph-based structural consistency regularization (SCR), which contains similarity and dissimilarity regularization terms by utilizing the positive and neglect edges, respectively, defined as

$$\sum_{i,j=1}^{N_S} W_{h+,ij}^{t1} \left\| Y_i' - Y_j' \right\|_2^2 + \sum_{i,j=1}^{N_S} W_{h-,ij}^{t1} f\left(\left\| Y_i' - Y_j' \right\|_2 \right) \tag{12}$$

where $f(\cdot)$ is a defined function.

In (12), the first term is equal to $2\text{Tr}(Y' \mathbf{L}_{h+}^{t1} Y'^{\mathrm{T}})$, i.e., the commonly used smoothness regularization for the unsigned graph. This term also corresponds to the first regularization term in the positive graph construction process of (4), which is utilized to constrain that the transformed $\tilde{Y}_i'$ and $\tilde{Y}_j'$ connected by $\mathcal{G}_{h+}^{t1}$ represent the same kind of objects.

In (12), the second term is used to penalize the similarity of the transformed $\tilde{Y}_i'$ and $\tilde{Y}_j'$ connected by $\mathcal{G}_{h-}^{t1}$, which prompts $\tilde{Y}_i'$ and $\tilde{Y}_j'$ to represent the different kinds of objects. Therefore, when designing the penalty function $f(d)$, it is necessary to consider that its purpose is not to increase the difference between $\tilde{Y}_i'$ and $\tilde{Y}_j'$ without a clear rationale. In other words, when $\tilde{Y}_i'$ and $\tilde{Y}_j'$ connected by $\mathcal{G}_{h-}^{t1}$ are very similar (indicating the same type of land cover), $f(d)$ should impose a large penalty on them; however, when $\tilde{Y}_i'$ and $\tilde{Y}_j'$ are not very similar (indicating they already belong to different land cover classes), $f(d)$ should no longer penalize them, otherwise it will introduce errors if $\tilde{Y}_i'$ and $\tilde{Y}_j'$ were further separated apart. So, the value of penalty function $-f(d)$ should be large around $d \approx 0$, and relatively stable and approaching 0 when d is relatively larger. For example, the following penalty function
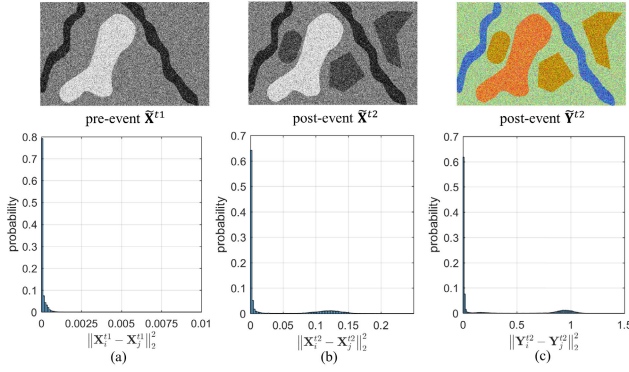
Fig. 4. Histogram of vertex feature distance on the graph $\mathcal{G}_{t1}^{h+}$ with different images: (a) pre-event image $\tilde{X}^{t1}$, (b) post-event image $\tilde{X}^{t2}$, and (c) post-event image $\tilde{Y}^{t2}$.

Fig. 5. Curves of the RGLR.

can be used as a reference:

$$f(d) = -\frac{1}{d^2 + \epsilon} \tag{13}$$

where $\epsilon > 0$ is a small parameter that tunes the gradient of the penalty function. In this article, we set ϵ being the average feature distance between superpixels $\tilde{Y}_i$ and $\tilde{Y}_j$, $(i, j) \in \mathcal{E}_{h+}^{t2}$ connected by the subgraph $\mathcal{G}_{h+}^{t2}$.

2) Distribution-Induced Reweighted Graph Laplacian Term: The signed graph-based SCR of (12) describes the structure connection between the transformed image and image from the source domain, that is, the structural consistency between $\tilde{Y}'$ and $\tilde{X}$. Next, we consider the structural associations between transformed images and images from the target domain, that is, $\tilde{Y}'$ and $\tilde{Y}$, $\tilde{X}'$ and $\tilde{X}$.

In the signal decomposition model, it decomposes the original feature matrices $(\mathbf{X}, \mathbf{Y})$ into transformed feature matrices $(\mathbf{X}', \mathbf{Y}')$ and changed feature matrices (Δ^x, Δ^y), denoted as $\mathbf{X} = \mathbf{X}' - \Delta^x$ and $\mathbf{Y} = \mathbf{Y}' - \Delta^y$. To explore the structural associations/differences between the transformed images and images from the target domain in a more rigorous mathematical manner, we further seek the statistical descriptions of the feature distance between vertices. As an illustrative example, Fig. 4 shows one pre-event image and two post-event image, denoted as $\tilde{X}^{t1}$, $\tilde{X}^{t2}$, and $\tilde{Y}^{t2}$, respectively, where $\tilde{X}^{t1}$ and $\tilde{X}^{t2}$ are homogeneous images from the same domain. We construct a high-order probabilistic graph $\mathcal{G}_{t1}^{h+}$ for pre-event image and then examine the vertex feature distance distributions of different images on this graph $\mathcal{G}_{t1}^{h+}$.

The histogram of vertex feature distance on different images is shown in Fig. 4, where the x-axis is the feature distances of $\|\mathbf{X}_i^{t1} - \mathbf{X}_j^{t1}\|_2^2$, $\|\mathbf{X}_i^{t2} - \mathbf{X}_j^{t2}\|_2^2$, and $\|\mathbf{Y}_i^{t2} - \mathbf{Y}_j^{t2}\|_2^2$ with $(i, j) \in \mathcal{E}_{h+}^{t1}$, and the y-axis is the fractions of feature distances for different vertices. From Fig. 4, we can make the following important observations.

Both the post-event images of $\tilde{X}^{t2}$ and $\tilde{Y}^{t2}$ have bimodal distributions of vertex feature distances in the similarity graph $\mathcal{G}_{t1}^{h+}$ constructed on pre-event image $\tilde{X}^{t1}$.

In the backward transformation process (11), it aims to translate the post-event image to the domain of the pre-event image to obtain the transformed image. Therefore, in Fig. 4, the post-event $\tilde{X}^{t2}$ can be regarded as the ideal transformed
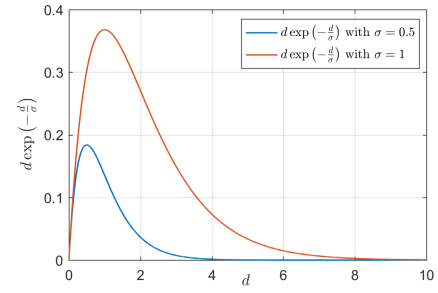
image of the output of the backward transformation. Consequently, we can obtain the following conclusion.

The transformed image should have a bimodal distribution of vertex feature distances in the similarity graph constructed on the image from the target domain.

The bimodal distribution property implies that the feature distance between vertices is either very large or very small. This property can be obtained by qualitative analysis in addition to being visualized through Fig. 4. For the unchanged vertices in the graph $\mathcal{G}_{h+}^{t1}$, based on the nature of structural consistency and similarity graph, it requires that the connected vertices $\tilde{X}_i'$ and $\tilde{X}_j'$ in transformed image should be also very similar, leading to a small feature distance. However, when one of the vertices changes and the other does not, we have that $\tilde{X}_i'$ and $\tilde{X}_j'$ in the transformed image should be very dissimilar, leading to a large feature distance. In other words, it is the change that causes this bimodal distribution.

To exploit this bimodal distribution property, we add a distribution-induced RGLR to the transformation model, defined as

$$
\begin{aligned}
\|\mathbf{X}'\|_{\text{RGL}} &= 2\text{Tr}\big(\mathbf{X}'\mathcal{L}(\mathbf{X}')\mathbf{X}'^{\text{T}}\big) \\
&= \sum_{i,j=1}^{N_S} \mathcal{W}\big(\mathbf{X}_i', \mathbf{X}_j'\big)\big\|\mathbf{X}_i' - \mathbf{X}_j'\big\|_2^2
\end{aligned}
\tag{14}
$$

where $\mathcal{L}(\mathbf{X}')$ is a functionalized Laplacian matrix of $\mathbf{X}'$ and $\mathcal{W}(\mathbf{X}_i', \mathbf{X}_j')$ is defined as

$$\mathcal{W}\big(\mathbf{X}_i', \mathbf{X}_j'\big) = W_{h+,ij}^{t1} \exp\big(-\big\|\mathbf{X}_i' - \mathbf{X}_j'\big\|_2^2/\sigma_{ij}^x\big). \tag{15}$$

Then, for the vertices connected by $\mathcal{G}_{h+}^{t1}$, the regularizer can be rewritten as $\mathcal{W}(\mathbf{X}_i', \mathbf{X}_j')\|\mathbf{X}_i' - \mathbf{X}_j'\|_2^2 = W_{h+,ij}^{t1}\text{dist}_{ij}^{x'} \exp(-\text{dist}_{ij}^{x'}/\sigma_{ij}^x)$ with $\text{dist}_{ij}^{x'} = \|\mathbf{X}_i' - \mathbf{X}_j'\|_2^2$, which has two minima at $\text{dist}_{ij}^{x'} = 0$ and $\text{dist}_{ij}^{x'} \to +\infty$, and one maximum at $\text{dist}_{ij}^{x'} = \sigma_{ij}^x$, as shown in the curve of Fig. 5. Therefore, we can find that this reweighted graph Laplacian term $\|\mathbf{X}'\|_{\text{RGL}}$ (14) prompts the backward transformed image $\tilde{X}'$ to follow a bimodal distribution of vertex feature distances in the similarity graph $\mathcal{G}_{h+}^{t1}$ constructed on pre-event image, which pushes for smaller values of $\text{dist}_{ij}^{x'}$ when $\text{dist}_{ij}^{x'} < \sigma_{ij}^x$ and larger values of $\text{dist}_{ij}^{x'}$ when $\text{dist}_{ij}^{x'} > \sigma_{ij}^x$. In addition, σ_{ij}^x can be regarded as a judgment threshold. In this article, we set $\sigma_{ij}^x = \text{dist}_{i(k_{\max})}^x + \text{dist}_{j(k_{\max})}^x$, that is, the sum of the $k_{\max}$th smallest values of the feature distance vectors dist_i^x and dist_j^x of the original pre-event image.

Similarly, for the forward transformation (11), we can define the reweighted graph Laplacian term $\|\mathbf{Y}'\|_{\text{RGL}}$ as

$$\|\mathbf{Y}'\|_{\text{RGL}} = 2\text{Tr}\left(\mathbf{Y}'\mathcal{L}(\mathbf{Y}')\mathbf{Y}'^{\text{T}}\right)$$

$$= \sum_{i,j=1}^{N_S} \mathcal{W}(\mathbf{Y}'_i, \mathbf{Y}'_j)\|\mathbf{Y}'_i - \mathbf{Y}'_j\|_2^2. \quad (16)$$

This $\|\mathbf{Y}'\|_{\text{RGL}}$ can prompt the transformed image $\tilde{Y}'$ to follow a bimodal distribution of vertex feature distances in the subgraph $\mathcal{G}_{h+}^{t2}$ built on the post-event image.

3) Change Sparsity Term: Similar to the previous transformation methods [33], we introduce a change sparsity regularization (CSR) that constrains the change image to be sparse, i.e., only a small number of pixel values are nonzero. Intuitively, the $\ell_{2,0}$-norm regularization, i.e., $\|\Delta^x\|_{2,0}$ and $\|\Delta^y\|_{2,0}$ that equals the number of changed superpixels, should be used as the CSR. However, given that the nonconvexity of $\ell_{2,0}$-norm will make it difficult to solve the optimization model, we use the convex relaxation forms of $\|\Delta^x\|_{2,1} = \sum_{i=1}^{N_S}\|\Delta_i^x\|_2$ and $\|\Delta^y\|_{2,1} = \sum_{i=1}^{N_S}\|\Delta_i^y\|_2$ instead.

4) Objective Function and Change Extraction: By incorporating the SCR (12), RGLR (16) and CSR, we have the signed graph model-based forward image transformation as

$$\min_{\mathbf{Y}',\Delta^y} \ 2\text{Tr}\left(\mathbf{Y}'\mathbf{L}_{h+}^{t1}\mathbf{Y}'^{\text{T}}\right) + \alpha \sum_{i,j=1}^{N_S} W_{h-,ij}^{t1} f\left(\|\mathbf{Y}'_i - \mathbf{Y}'_j\|_2\right)$$

$$+ \beta\|\mathbf{Y}'\|_{\text{RGL}} + \lambda\|\Delta^y\|_{2,1}$$

$$\text{s.t.} \quad \mathbf{Y}' = \mathbf{Y} + \Delta^y \quad (17)$$

where $\alpha, \beta, \lambda > 0$ denote the weighting parameters.

It could be seen from (17) that the first term of similarity regularization prefers to output smooth $\mathbf{Y}'$ on the graph $\mathcal{G}_{h+}^{t1}$, such as $\mathbf{Y}' = 0$, the second term of dissimilarity regularization prefers to output nonsmooth $\mathbf{Y}'$ on the graph $\mathcal{G}_{h-}^{t1}$, the third term of distribution-induced RGLR prefers to output bimodal $\mathbf{Y}'$ on the graph $\mathcal{G}_{h+}^{t2}$, the last term of change sparsity regularization prefers to output the unchanged $\mathbf{Y}'$ of $\mathbf{Y}' = \mathbf{Y}$. Therefore, we can observe that the regularization terms in (17) are antagonistic to each other, which can help the model output stable results.

Similarly, for the backward transformation process, we have the following model:

$$\min_{\mathbf{X}',\Delta^x} \ 2\text{Tr}\left(\mathbf{X}'\mathbf{L}_{h+}^{t2}\mathbf{X}'^{\text{T}}\right) + \alpha \sum_{i,j=1}^{N_S} W_{h-,ij}^{t2} f\left(\|\mathbf{X}'_i - \mathbf{X}'_j\|_2\right)$$

$$+ \beta\|\mathbf{X}'\|_{\text{RGL}} + \lambda\|\Delta^x\|_{2,1}$$

$$\text{s.t.} \quad \mathbf{X}' = \mathbf{X} + \Delta^x. \quad (18)$$

The minimization problems of (17) and (18) can be approximately solved by the alternating direction method of multipliers (ADMM), which is detailed in Appendix A.

Once the transformed feature matrices and changed feature matrices are obtained from the transformation models of (17) and (18), we can calculate the transformed images as $\tilde{X}' = \mathcal{F}^{-1}(\mathbf{X}')$, $\tilde{Y}' = \mathcal{F}^{-1}(\mathbf{Y}')$, and compute the forward and backward DIs as

$$\left.\begin{array}{l}\text{DI}^y(m,n) = \left\|\Delta_i^y\right\|_2 \\ \text{DI}^x(m,n) = \left\|\Delta_i^x\right\|_2\end{array}\right\} \ (m,n) \in \Lambda_i, \quad i = 1,\ldots,N_S \quad (19)$$

where $\|\Delta_i^y\|_2$ and $\|\Delta_i^x\|_2$ indicate the change level of the ith superpixel in forward and backward transformation models, respectively. Since the forward and backward DIs both describe the same change event, they can be fused to improve change detection accuracy. We employ the MRF co-segmentation [26] to calculate the final CM, which integrates DI fusion and DI segmentation into a unified model to simultaneously improve the fusion and segmentation effects, yielding more accurate binary CM. The framework of SGIT is outlined in Algorithm 1.

Algorithm 1 SGIT-Based HeCD

Input: Heterogeneous images of $\tilde{X}$, $\tilde{Y}$, parameters of N_S, α, β, and λ.
Preprocessing:
    Segment multitemporal images into superpixels.
    Extract superpixel features to compose $\mathbf{X}$ and $\mathbf{Y}$.
Structure representation:
    Construct the subgraphs of $\mathcal{G}_+^{t1}$ and $\mathcal{G}_+^{t2}$ using (5).
    Construct the subgraphs of $\mathcal{G}_-^{t1}$ and $\mathcal{G}_-^{t2}$ by negative sampling.
    Construct the high-order signed graphs of $\mathcal{G}_{h\pm}^{t1}$ and $\mathcal{G}_{h\pm}^{t2}$.
Image transformation:
    Initialize: set Δ^x, $\Delta^y = \mathbf{0}$.
    Repeat:
        1: Update $\mathbf{X}'$ and $\mathbf{Y}'$ using (25).
        2: Update Δ^x and Δ^y using (28).
        3: Update $\mathbf{R}$ using (29).
    Until the stopping criterion is met.
Change extraction:
    Calculate the DI^x and DI^y using (19).
    Calculate the CM through MRF co-segmentation.

IV. EXPERIMENTAL RESULTS AND DISCUSSIONS

A. Experimental Setting

1) Datasets: To test the performance of the proposed algorithm, we employ five widely used real HeCD datasets listed in Table I, where the pre- and post-event images are sourced from different optical sensors (i.e., Datasets #1–#3) or different types of sensors (i.e., Datasets #4 and #5). In addition, these datasets cover different image sizes, different spatial resolutions, and different change events as shown in Fig. 6(a)–(c), which can effectively verify the detection performance of different methods in different HeCD task scenarios.

2) Evaluation Metrics: We employ two types of metrics to assess DI and CM, respectively. For DI, we plot its receiver operating characteristic (ROC) curve and precision–recall (PR) curve and calculate the corresponding areas under the ROC and PR curves, denoted as AUR and AUP, respectively. For CM, we compute its confusion matrix to obtain the true positive (TP), false positive (FP), true negative (TN), and false negative (FN), from which we calculate the overall accuracy (OA), $F1$-score, and Kappa coefficient (κ).

3) Parameter Setting: For all the experiment results of the proposed SGIT, we set the superpixel number $N_S \approx 2500$, and fix the weighting parameters as $\beta = 4$, $\lambda = 0.1$, and adjust

TABLE I
INTRODUCTION TO THE HETEROGENEOUS DATASETS

Dataset	Sensor (or source)	Date	Image size (pixels)	Location	Spatial resolution	Change event
#1	Landsat-5 Google Earth	Sept. 1995 July 1996	$300 \times 412 \times 1$ $300 \times 412 \times 3$	Sardinia, Italy	30m	Lake expansion
#2	Pleiades WorldView2	May 2012 July 2013	$2000 \times 2000 \times 3$ $2000 \times 2000 \times 3$	Toulouse, France	0.52m	Construction
#3	Landsat-5 EO-1 ALI	Aug. 2011 Sept. 2011	$1534 \times 808 \times 7$ $1534 \times 808 \times 10$	Texas, USA	30m	Forest fire
#4	Radarsat-2 Google Earth	June 2008 Sept. 2012	$593 \times 921 \times 1$ $593 \times 921 \times 3$	Shuguang Village, China	8m	Farmland changes
#5	QuickBird-2 TerraSAR-X	July 2006 July 2007	$4135 \times 2325 \times 3$ $4135 \times 2325 \times 1$	Gloucester, England	0.65m	Flooding

the α with $\alpha \in \{2^0, 2^1, 2^2, 2^3, 2^4\}$. These parameters will be discussed in Section IV-C.

B. Experimental Results

1) Difference Images: To fully assess the DI computed by SGIT in distinguishing changes, we chose five state-of-the-art (SOTA) comparison HeCD methods, including the AMD base image regression (AMDIR) [19], sparse constrained adaptive structural consistency-based method (SCASC) [33], conditional generative adversarial network (CGAN) [8], the adversarial cyclic encoder network (ACE-Net) [22], and the similarity- and dissimilarity-based multimodal change detection method (SDMCD) [34]. Furthermore, for SCASC and CGAN, which involve only unidirectional transformation processes, we obtain DI for the other direction by swapping the order of input images.

Fig. 6 shows DIs computed by different methods on the forward and backward transformations, from which significant performance differences can be noted between the forward and backward DIs from the same method. Some methods achieve good performance in one direction, obtaining high-quality DI, while completely failing in the other direction, resulting in poor DI performance. For example, SCASC can detect changes in forward transformation on Datasets #1, #3, and #5, but cannot distinguish changes in backward transformation. In addition, ADMIR exhibits similar behavior on Datasets #1 and #4. Conversely, the performance of backward DIs of CGAN and ACE-Net on Dataset #5 is much better than that of forward DIs. This also validates two aspects. First, the challenge of forward and backward transformations is different. For example, transformations from complex to simple scenes and from simple to complex scenes, as well as transformations from optical to SAR images and from SAR to optical images, carry different difficulties. Second, it is necessary to fuse forward and backward DIs, as they both reflect the same change event and a reasonable fusion will improve change detection accuracy.

At the same time, it is also found from Fig. 6 that SDMCD and the proposed SGIT are able to achieve relatively consistent detection results in both forward and backward transformations. This is due to the fact that compared to AMDIR, SCASC, and ACE-Net, which only consider the similarity relationship, SDMCD considers the dissimilarity relationship within the image, and constructs a KFN graph to capture the
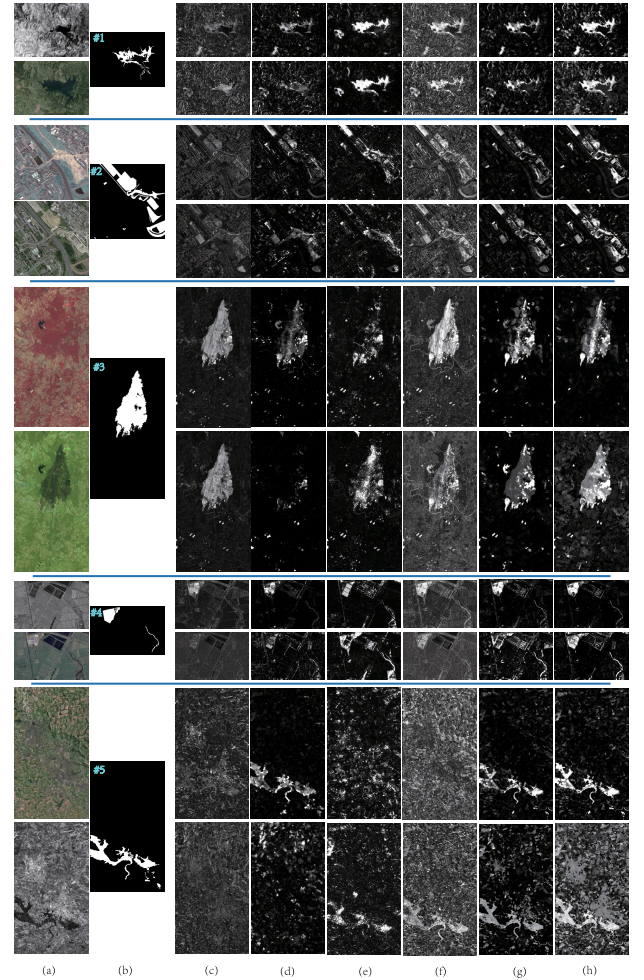
Fig. 6. Heterogeneous datasets and DIs. From top to bottom, they are corresponding to Datasets #1–#5, respectively. (a) Pre- and post-event images. (b) Ground truth. From (c)–(h) are the forward DI^y and backward DI^x obtained by: (c) AMDIR, (d) SCASC, (e) CGAN, (f) ACE-Net, (g) SDMCD, and (h) proposed SGIT, respectively.

high-frequency image structural features as a complement to the KNN graph that characterizes the low-frequency structural features. Different from SDMCD, the SGIT proposed in this article starts from the difference between unsigned and signed graphs under the perspective of the GSP, and then constructs a signed graph model that describes the "attraction" and "repulsion" inside the image, which can more fully characterize the
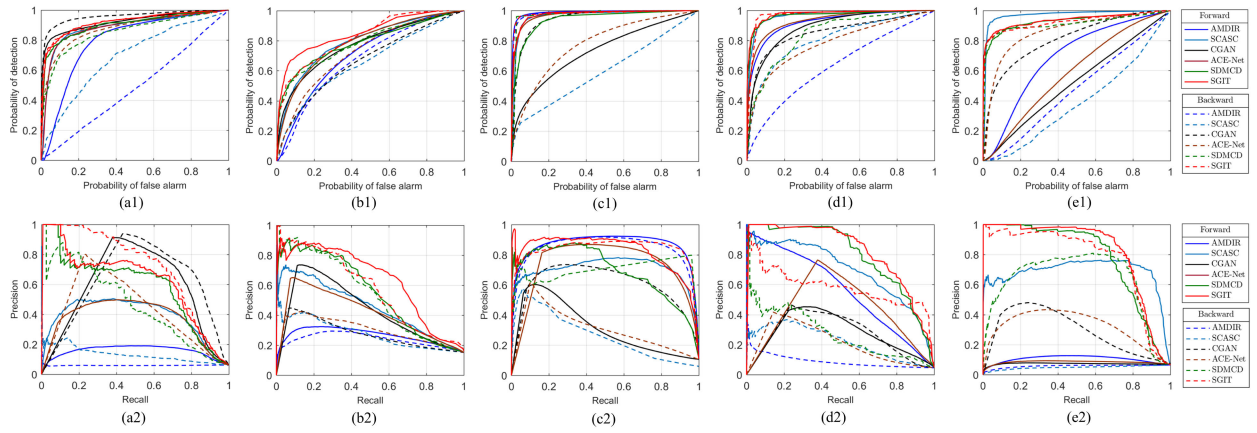
Fig. 7. ROC curves (top row) and PR curves (bottom row) of DIs on different datasets. (a1) ROC: Dataset #1. (a2) PR: Dataset #1. (b1) ROC: Dataset #2. (b2) PR: Dataset #2. (c1) ROC: Dataset #3. (c2) PR: Dataset #3. (d1) ROC: Dataset #4. (d2) PR: Dataset #4. (e1) ROC: Dataset #5. (e2) PR: Dataset #5.

TABLE II

AUR AND AUP OF DIs. THE HIGHEST SCORES ARE HIGHLIGHTED IN BOLD

Methods	Dataset #1				Dataset #2				Dataset #3				Dataset #4				Dataset #5				Average			
	Forward DI		Backward DI		Forward DI		Backward DI		Forward DI		Backward DI		Forward DI		Backward DI		Forward DI		Backward DI		Forward DI		Backward DI	
	AUR	AUP	AUR	AUP	AUR	AUP	AUR	AUP	AUR	AUP	AUR	AUP	AUR	AUP	AUR	AUP	AUR	AUP	AUR	AUP	AUR	AUP	AUR	AUP
AMDIR [19]	0.7953	0.1550	0.4823	0.0598	0.7082	0.2636	0.6755	0.2365	**0.9860**	**0.8564**	**0.9819**	**0.8332**	0.9218	0.5643	0.6308	0.0898	0.6907	0.1033	0.4751	0.0570	0.8204	0.3885	0.6491	0.2553
SCASC [33]	0.8846	0.3831	0.6883	0.1325	0.7931	0.4575	0.6523	0.2760	0.9637	0.6990	0.5638	0.2689	0.9676	0.6947	0.8062	0.2181	**0.9711**	0.6814	0.3883	0.0477	0.9160	0.5831	0.6198	0.1887
CGAN [8]	**0.9225**	0.5505	**0.9572**	0.5896	0.7729	0.4362	0.6671	0.2842	0.7303	0.3085	0.9446	0.5968	0.8901	0.2734	0.8525	0.2701	0.5214	0.0726	0.8118	0.2865	0.7675	0.3282	0.8466	0.4055
ACE-Net [22]	0.8865	0.3679	0.8769	0.4175	0.7787	0.4120	0.7193	0.3071	0.9695	0.7187	0.7768	0.3011	0.9336	0.4197	0.7824	0.2039	0.5887	0.0832	0.8594	0.3184	0.8314	0.4003	0.8030	0.3096
SDMCD [34]	0.9018	0.5850	0.8526	0.4681	0.7871	0.5536	0.7844	0.5432	0.9399	0.6906	0.9718	0.6932	0.9679	0.7911	0.8438	0.2740	0.9288	0.7836	0.9114	0.6178	0.9051	0.6808	0.8728	0.5193
proposed SGIT	0.9195	**0.6244**	0.9174	**0.7143**	**0.8544**	**0.6379**	**0.8266**	**0.5728**	0.9739	0.8229	0.9781	0.8289	**0.9770**	**0.8165**	**0.9698**	**0.5808**	0.9352	**0.8248**	**0.9251**	**0.7734**	**0.9320**	**0.7453**	**0.9234**	**0.6940**

image structural features. In turn, it can overcome the detection inconsistency problem caused by the asymmetric representation of image structural information caused by unsigned graph models in the pre- and post-event images. Furthermore, in contrast to other methods, SGIT more fully exploits the change prior information inherent in the HeCD problem and introduces a bimodal distribution-induced regularization in the transformation model, which improves the accuracy and robustness of change detection, as verified by DIs obtained by SGIT in Fig. 6(h).

Fig. 7 plots the ROC and PR curves of DIs generated by different methods in Fig. 6, where the corresponding AUR and AUP are presented in Table II. It can be found that the DIs generated by SGIT can well discriminate between changed and unchanged. SGIT gains the average AUR and AUP values of 0.9320 and 0.7453 on forward DI, respectively, which are 1.60% and 6.45% higher than those of the second-ranked (SCASC and SDMCD), respectively. At the same time, the average AUR and AUP values for the backward DI of SGIT are 0.9234 and 0.6940, which are 5.06% and 17.47% higher than those of the second-ranked SDMCD, respectively. From Figs. 6 and 7 and Table II, it can be noted that the DI obtained by SGIT provides a high capability of distinguishing between changed/unchanged, which in turn leads to a promising CM using simple threshold segmentation or clustering methods.

2) Change Maps: To more fully assess the ability of SGIT to detect changes, we compared the CMs produced by different methods, including not only the five methods of AMDIR, SCASC, CGAN, ACE-Net, and SDMCD in the DI comparison but also another five HeCD methods of the fractal projection

and MRF segmentation-based method (FPMS) [18], the circular invariant convolution model (CICM) [5], the unsupervised spatial self-similarity difference-based method (USSD) [13], two fully convolutional networks-based X-Net [22], the code-aligned autoencoders-based method (CAAE) [21]. For a more adequate comparison, for those methods that only consider one-way transformation processes in the original papers (such as FPMS, CGAN, and SCASC), we use the best results in their forward and backward transformation processes for comparison.

Fig. 8 shows the CMs of different methods on Datasets #1–#5, where the TP, FP, TN, and FN are marked with different colors for better visual comparison. Table III reports the OA, κ, and $F1$ of these CMs. Specifically, Dataset #1 reflects changes in lake expansion, with challenges including "pseudo-changes" caused by land areas and the loss of details in lake change contours. Visually, AMDIR, FPMS, and CAAE exhibit more false alarms, while SDMCD and CICM lose many change details, resulting in a high miss detection. In comparison, SGIT, USSD, CGAN, and X-Net detect more change areas with relatively fewer false alarms. Dataset #2 has a higher image resolution (0.52 m), richer image details and textures, and includes more land cover categories (e.g., roads, houses, trees, ball fields, construction sites, etc.) with an imbalanced percentage. The change scenarios in this dataset are more complex, resulting in low accuracy for most methods. For example, AMD and USSD exhibit a large number of false alarms, while SCASCC, CGAN, and CICM miss a lot of detections. Nevertheless, the proposed SGIT still achieves satisfactory results on this dataset, with an $F1$ score reaching 0.6868. Dataset #3 reflects changes in forests before and after
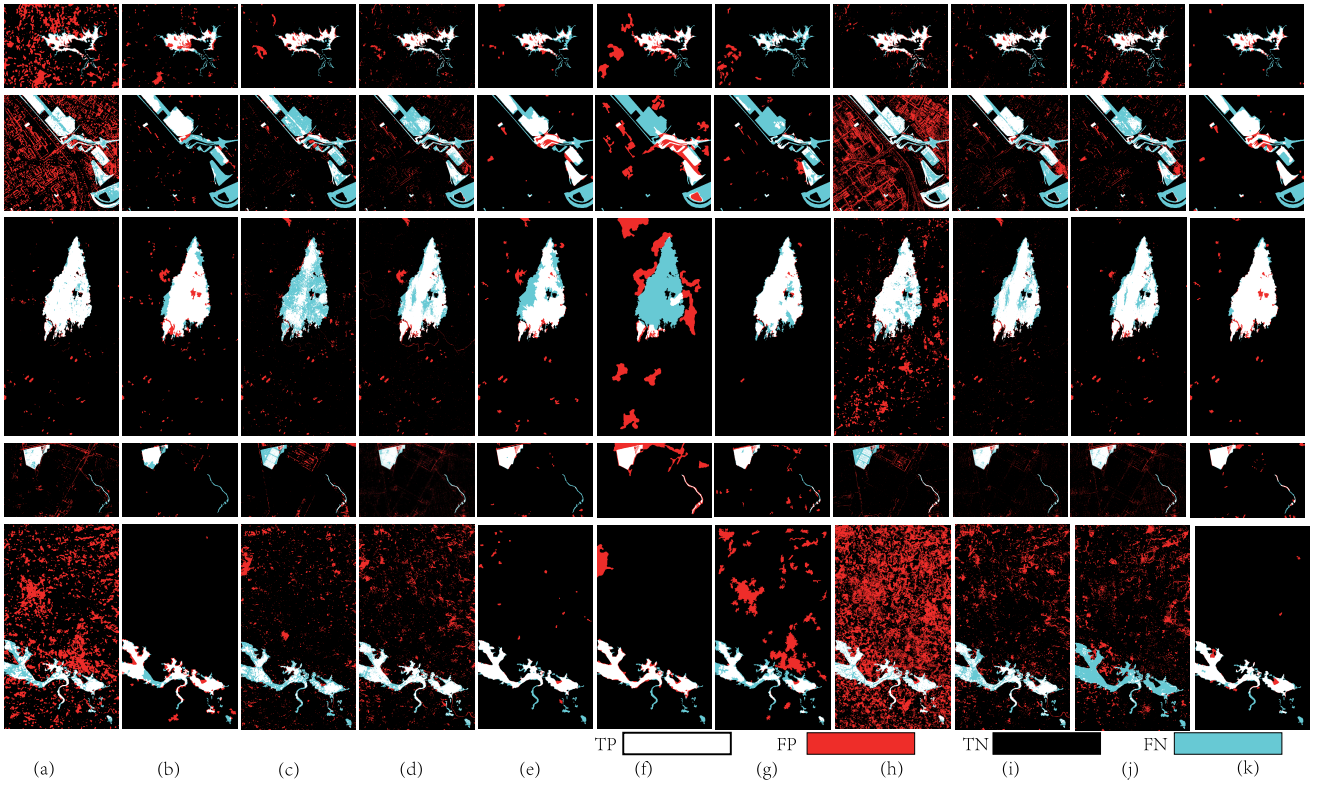
Fig. 8. CMs generated by different methods. From top to bottom, they are corresponding to Datasets #1–#5, respectively. From (a) to (k) are the CMs generated by: (a) AMDIR, (b) SCASC, (c) CGAN, (d) ACE-Net, (e) SDMCD, (f) FPMS, (g) CICM, (h) USSD, (i) X-Net, (j) CAAE, and (k) proposed SGIT, respectively. In the CM, white, red, black, and cyan mark the TP, FP, TN, and FN, respectively.

TABLE III
OA, κ AND $F1$ OF CMS. THE BEST SCORES ARE MARKED IN RED

Methods	Dataset #1			Dataset #2			Dataset #3			Dataset #4			Dataset #5			Average		
	OA	κ	F1	OA	κ	F1	OA	κ	F1	OA	κ	F1	OA	κ	F1	OA	κ	F1
AMDIR [19]	0.7994	0.2551	0.3280	0.7244	0.2588	0.4106	0.9845	0.9188	0.9275	0.9502	0.5722	0.5969	0.7821	0.1709	0.2538	0.8481	0.4352	0.5034
SCASC [33]	0.9468	0.5933	0.6214	0.8917	0.4638	0.5163	0.9703	0.8453	0.8620	0.9792	0.7407	0.7514	0.9725	0.7738	0.7885	0.9521	0.6834	0.7079
CGAN [8]	0.9653	0.7235	0.7420	0.8629	0.3382	0.4025	0.9204	0.4819	0.5216	0.9426	0.4020	0.4321	0.9110	0.3645	0.4123	0.9204	0.4620	0.5021
ACE-Net [22]	0.9584	0.6740	0.6961	0.8643	0.3893	0.4605	0.9608	0.7896	0.8114	0.9539	0.5858	0.6091	0.9009	0.4469	0.4968	0.9277	0.5771	0.6148
SDMCD [34]	0.9642	0.6716	0.6905	0.9189	0.6381	0.6828	0.9540	0.7434	0.7688	0.9823	0.8102	0.8194	0.9820	0.8432	0.8528	0.9603	0.7413	0.7629
FPMS [18]	0.9385	0.5932	0.6245	0.8265	0.2688	0.3681	0.8300	0.0195	0.1122	0.9382	0.5692	0.5970	0.9703	0.7705	0.7863	0.9007	0.4443	0.4976
CICM [5]	0.9425	0.4509	0.4809	0.8666	0.2702	0.3208	0.9834	0.9091	0.9183	0.9739	0.7453	0.7588	0.8960	0.3711	0.4230	0.9325	0.5493	0.5804
USSD [13]	0.9645	0.7257	0.7445	0.6939	0.2144	0.3793	0.9161	0.6317	0.6781	0.9422	0.3756	0.4059	0.5780	0.0837	0.1865	0.8189	0.4062	0.4789
X-Net [22]	0.9671	0.7226	0.7401	0.8683	0.4626	0.5382	0.9661	0.8147	0.8335	0.9639	0.6513	0.6699	0.8777	0.3407	0.4014	0.9286	0.5984	0.6366
CAAE [21]	0.9301	0.5335	0.5692	0.8609	0.3731	0.4459	0.9745	0.8606	0.8748	0.9432	0.5403	0.5679	0.8444	0.0136	0.0688	0.9106	0.4588	0.5053
proposed SGIT	0.9708	0.7386	0.7542	0.9197	0.6425	0.6868	0.9839	0.9179	0.9269	0.9824	0.8174	0.8267	0.9833	0.8545	0.8633	0.9680	0.7942	0.8116

wildfires. With the exception of CGAN and FPMS, which exhibit a high false-negative rate, most methods could detect change areas well. The challenges of Dataset #4 lie in detecting changes in farmland and river contours. From Fig. 8, it can be seen that some methods ignore changes in the river at the bottom right corner, such as AMDIR, SCASC, SDMCD, and X-Net. Dataset #5 reflects changes in rivers, whose challenges include wide image coverage, complex texture features of objects, and susceptibility to pseudo-changes. As shown in Fig. 8, SCASC, ACE-Net, SDMCD, FPMS, and SGIT are all able to detect changes comprehensively, but there are also a few errors.

Generally, from both Fig. 8 and Table III, the proposed SGIT not only detects more complete change areas but also exhibits relatively low false-negative rates. The average κ and $F1$ values of SGIT are 0.7942 and 0.8116, respectively,

which are higher than those of the second-ranked SDMCD at 5.29%, and 4.87%, respectively. This can be attributed to two factors: first, SGIT utilizes a high-order unsigned graph model for a more comprehensive representation of image structures; second, it employs bimodal distribution-induced regularization in the image transformation model for more effective utilization of change prior information.

In addition, apart from the ten comparison methods in Fig. 8 and Table III, we also list the κ values obtained by some other SOTA methods in Table IV, where we directly quote the results from their original published papers for the sake of fairness. Due to the different datasets used in each paper, Table IV is not aligned. Among these approaches, COMIC [20], C3D [16], DADR [17], SDCGAE [39], MDCTNet [23], IOECL [9], and GCDG [32] are deep learning-based methods. As can be seen from Table V, the proposed SGIT consistently achieves
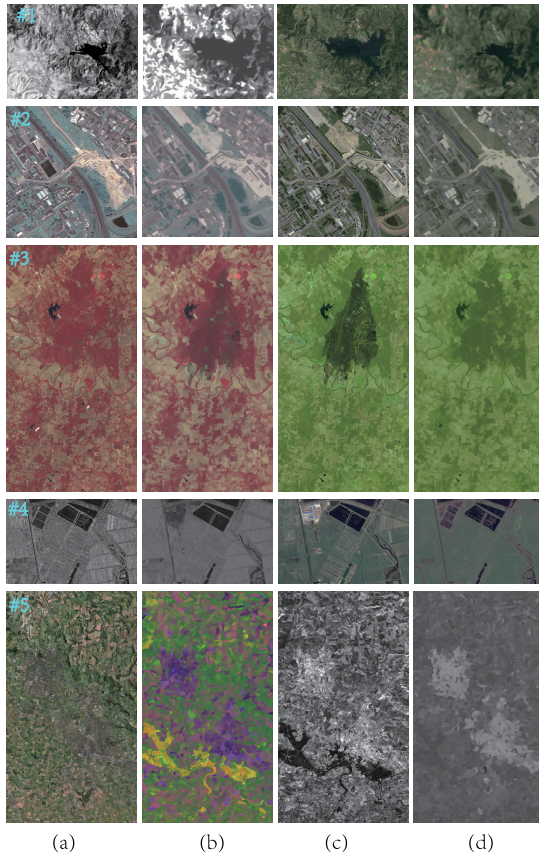
Fig. 9. Transformed images of SGIT on different datasets. (a) Original pre-event $\tilde{X}$. (b) Backward transformed $\tilde{X}'$. (c) Original post-event $\tilde{Y}$. (d) Forward transformed $\tilde{Y}'$.

Fig. 10. Sensitivity analysis of parameter N_S.

better or highly competitive κ across various datasets, further demonstrating the effectiveness of the signed graph-based HeCD method.

C. Discussions

1) Similarity Between Transformed and Original Images: Fig. 9 shows the transformed images obtained by SGIT on Datasets #1–#5. Comparing the forward/backward transformed images with the original pre-event/post-event images in Fig. 9, it can be found that: first, the internal structural relationships between the transformed images and the original source domain images are consistent, i.e., $\tilde{X}'$ and $\tilde{Y}$, $\tilde{Y}'$ and $\tilde{X}$ share the similarity/dissimilarity relationships, respectively. Second, the style of the transformed images is consistent with that of the original images from target domain, i.e., the image pixel values of $\tilde{X}'$ and $\tilde{X}$, $\tilde{Y}'$, and $\tilde{Y}$ are comparable, demonstrating that SGIT can achieve cross-domain image transformation. Third, the differences between the transformed images and the original target domain images reflect change information, indicating that SGIT can utilize the image structure to accomplish change detection between heterogeneous images.

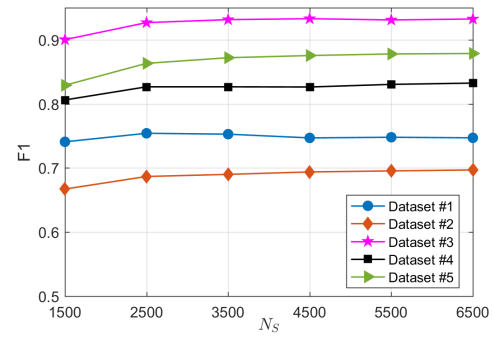*2) Parameters Analysis:* The main parameters used in SGIT are the superpixel number of N_S and the weighting parameters of α and β.

For the parameter N_S, it should be determined according to the image spatial resolution and timeliness requirement of the change detection task. We vary N_S from 1500 to 6500 with steps of 1000, and plot the corresponding $F1$ values obtained by SGIT on different datasets in Fig. 10. Generally, it can be found that the $F1$ initially rises significantly as N_S increases from 1500 to 2500, rises slowly as N_S increases from 2500 to 4500, and stabilizes as N_S increases from 4500 to 6500. When the N_S is too small, it will result in the segmented superpixel containing different land cover types internally, i.e., compromising the homogeneity within the superpixels. This will make it difficult to extract superpixel features and lead to inaccurate characterization of the image structure, which in turn affects the performance of the image structure-based SGIT. On the contrary, when the value of N_S increases, although it improves change detection accuracy, it also increases the size of the graph model, resulting in increased computational complexity and affecting algorithm efficiency. As shown in Fig. 10, SGIT is relatively robust to N_S, which can be attributed to two factors: First, SGIT is a topological structure-based method that focuses on the overall structure of the image (i.e., the relationships between superpixels) rather than the internal features of individual superpixels. Second, SGIT employs strategies such as adaptive weighting, negative sampling and neighborhood expansion, which enhance the representational capability and robustness of the graph model, reducing its sensitivity to the parameter N_S. Comprehensively considering the accuracy and timeliness requirements of the HeCD task, we set $N_S = 2500$ as a compromise in SGIT.

For the weighting parameters of α and β, they are employed for balancing different regularization in the image transformation model (17). Fig. 11 shows the performance of SGIT with varying α and β (from 2^{-1} to 2^5 with ratio 2). It can be found that both forward and backward image transformation models are relatively robust to the parameter β, which is suggested to take a reference range from 2^0 to 2^3 (we fix $\beta = 4$ in this article). As can be seen from Fig. 11, the forward and backward image transformations of SGIT require different values for the α. Taking Dataset #3 as an example, SGIT needs a larger α in the backward transformation and a smaller α in the forward transformation. This is because in Dataset #3, the pre-event image and post-event image exhibit the same similarity relationships on the unsigned subgraph

TABLE IV

κ OF CMS GENERATED BY DIFFERENT METHODS, WHICH ARE DIRECTLY QUOTED FROM THEIR ORIGINAL PUBLISHED PAPERS

| Dataset #1	κ	Dataset #2	κ	Dataset #3	κ	Dataset #4	κ	Dataset #5	κ
C3D [16]	0.799	**SGIT**	**0.642**	**SGIT**	**0.918**	Dataset #4	κ	Dataset #5	κ
COMIC [20]	0.761	IOECL [9]	0.553	MDCTNet [23]	0.902	C3D [16]	0.877	C3D [16]	0.912
SGIT	**0.739**	AOSG [27]	0.513	FDMCD [25]	0.880	**SGIT**	**0.817**	**SGIT**	**0.854**
DADR [17]	0.739	EGSR [29]	0.503	COMIC [20]	0.878	MDCTNet [23]	0.810	SDCGAE [39]	0.795
SDCGAE [39]	0.719	SDCGAE [39]	0.493	DADR [17]	0.878	FDMCD [25]	0.804	AGSCC [24]	0.766
FDMCD [25]	0.714	AGSCC [24]	0.490	GBFCD [42]	0.888	AGSCC [24]	0.773	AOSG [27]	0.759
AGSCC [24]	0.658	MDCTNet [23]	0.416	KCCA [7]	0.650	GCDG [32]	0.750	GBFCD [42]	0.156
MDCTNet [23]	0.618	GCDG [32]	0.350						
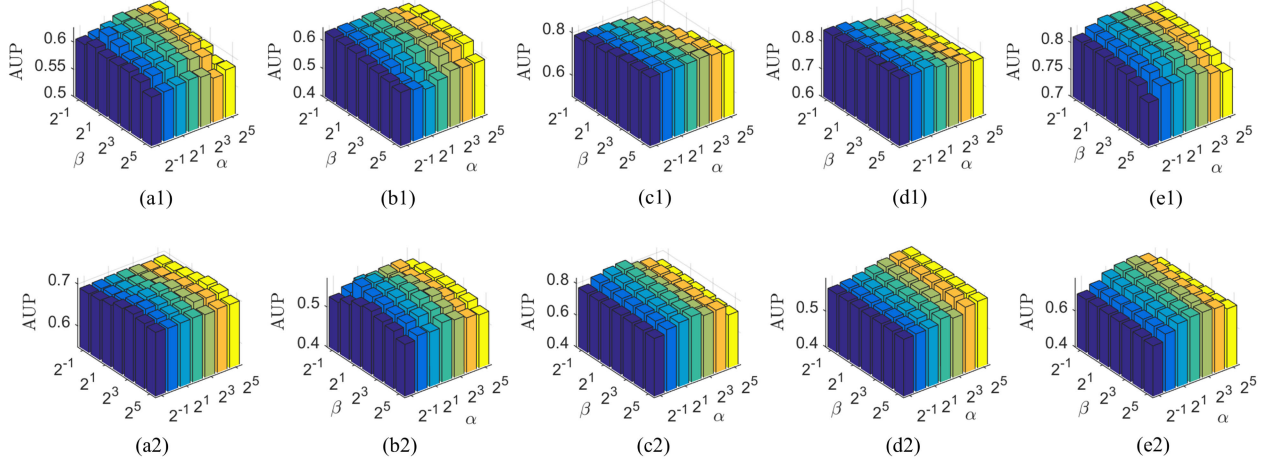
Fig. 11. Sensitivity analysis of parameters α and β. The top and bottom rows are the AUP for forward and backward DIs, respectively. (a1) Dataset #1: forward. (a2) Dataset #1: backward. (b1) Dataset #2: forward. (b2) Dataset #2: backward. (c1) Dataset #3: forward. (c2) Dataset #3: backward. (d1) Dataset #4: forward. (d2) Dataset #4: backward. (e1) Dataset #5: forward. (e2) Dataset #5: backward.

model $\mathcal{G}_+^{t2}$, but exhibit quite different similarity relationships on the unsigned subgraph model $\mathcal{G}_+^{t1}$. Therefore, in forward transformation, it is possible to detect part of the change region using only the unsigned graph model, such as the forward DI of SCASC in Fig. 6, whereas in the backward transformation, it is hard to detect changes using only the unsigned graph model. Consequently, a larger α is needed to enhance the role of negative edges of the signed graph model in the backward transformation process.

3) Ablation Study: SGIT mainly contains the structure representation based on the signed graph and image transformation based on the structural consistency. Compared to the previous image structure-based HeCD methods (such as SCASC, SDMCD), the key techniques of SGIT are: the exploitation of signed graph, i.e., the negative edge-induced penalty (NEP) in the transformation model (17), and the exploitation of change-induced bimodal distribution, i.e., the distribution-induced reweighted RGLR of (14). To analyze the role of NEP and RGLR in the proposed method, we build a baseline model: an image transformation model without using NEP and RGLR, which is similar to the previous SCASC of (1) and (2).

Table V reports the average quantitative results of Datasets #1–#5 from the image transformation models with/without NEP and RGLR. From the results in Table V, it can be found that the change detection performance of SGIT significantly

TABLE V

ABLATION STUDY OF SGIT, WHERE "BASE" REPRESENTS THE BASELINE

Settings	Forward DI		Backward DI		Final CM		
	AUR	AUP	AUR	AUP	OA	κ	F1
Base	0.9190	0.5891	0.6358	0.2146	0.9529	0.6907	0.7140
Base+RGLR	0.9231	0.6433	0.7417	0.3389	0.9582	0.7334	0.7528
Base+NEP	0.9287	0.7215	0.9001	0.6260	0.9611	0.7713	0.7902
Base+RGLR+NEP	0.9320	0.7453	0.9234	0.6940	0.9680	0.7942	0.8116

decreases when NEP and RGLR are not used. Specifically, by constructing the signed graph proposed in Section III-B for characterizing the image structure, which uses positive/negative edge weights to depict the attractive/repulsive forces between superpixels within an image, respectively, the image structure can be more adequately represented than using only unsigned graph model with positive edge weights, thereby enabling a more accurate image transformation model. As listed in Table V, when the NEP is employed, SGIT's detection performance is greatly improved, especially on the DI obtained from backward transformation. For example, the average AUP of backward DI is improved by almost 0.4114 compared to the baseline. This can also be verified from Fig. 6, where the backward transformation for Datasets #3 and #5 would entirely fail when only using the unsigned KNN graph with positive edge weight. By introducing the

change-induced bimodal distribution-based RGLR into the image transformation model, the structure of the transformed image fits the actual distribution better, leading to more accurate transformed images and DIs. To be specific, when incorporating the RGLR into the baseline model, the AUP of forward and backward DIs are improved by 0.0542 and 0.1243, respectively. Furthermore, when both NEP and RGLR are used, the change detection performance of SGIT is enhanced even more, with an average κ improvement of nearly 0.1035 over the baseline model.

4) Computational Complexity: The main computational burden of SGIT is concentrated in the structure representation and image transformation stages. For the former, it requires $\mathcal{O}((B_1 + B_2)N_S^2)$ and $\mathcal{O}(N_S^2 \log N_S)$ for computing and sorting the feature distance matrix, respectively, it requires $\mathcal{O}(N_S^2)$ for computing the positive weighting matrix and negative sampling probabilities, additionally, it also requires $\mathcal{O}(N_S^2)$ for computing the high-order positive and negative weighting matrices. For the latter, it requires $\mathcal{O}((B_1 + B_2)N_S^2)$ for solving the $\mathbf{X}'$ and $\mathbf{Y}'$-subproblems, it requires $\mathcal{O}((B_1 + B_2)N_S)$ for solving the Δ^x and Δ^y-subproblems, and it also requires $\mathcal{O}((B_1 + B_2)N_S)$ for updating the Lagrange multipliers. The computational times (in seconds) for traditional comparison methods on Dataset #2 (with image size of $2000 \times 2000 \times 3$) are as follows: AMDIR: 109.6; SCASC: 26.7; SDMCD: 49.1; FPMS: 92.4; CICM: 113.0; USSD: 795.2; and SGIT: 61.5. The C++ implementations of FPMS and CICM were executed on a Linux computer with an Intel Xeon Silver 4110 CPU, while AMDIR, SCASC, SDMCD, USSD, and SGIT were run in MATLAB 2016a on a Windows desktop equipped with an Intel Core i7-8700K CPU. It can be observed that SGIT demonstrates a moderate level of computational efficiency. In the future, we will consider employing more efficient optimization algorithms to solve the image transformation model, aiming to further enhance the detection efficiency of SGIT.

V. CONCLUSION

This article focuses on the issue of unsupervised HeCD of heterogeneous remote sensing images. We first analyze the limitations of previous methods based on unsigned graphs in addressing the HeCD problem and find that in some scenarios, using only unsigned graphs will lead to failure in change detection due to the limited capability in capturing image structures. Based on this analysis, this article proposes a signed graph-based image transformation method for HeCD, which mainly consists of two processes: structure representation and image transformation. First, it constructs a signed graph model, which utilizes positive/negative weights to portray the similarity/dissimilarity relationships within an image, respectively. Additionally, adaptive weights, negative sampling, and balance theory-based neighbor expansion strategies are employed to enhance the structure representation capability of the signed graph. Subsequently, a structural consistency-based image transformation model is proposed to decompose the target images into transformed and changed images, which requires the transformed image and original image to have the same structural property on the signed graph. Moreover, this article

reveals that changes can induce a bimodal distribution of vertex feature distances in the transformed images, which is then incorporated as a constraint into the image transformation model to further enhance the model's accuracy. Extensive comparative experiments validate the effectiveness of the proposed method across various HeCD scenarios. In the future, we will further analyze the properties of different signed graph models in the graph spectral domain, serving to guide the construction of signed graphs and propose more accurate structural constraints to improve the change detection performance.

APPENDIX A
GRAPH SIGNAL SMOOTHNESS FOR SIGNED GRAPH

In the GSP, for the graph signal $\mathbf{Y}'$ and signed graph with weight $\mathbf{W}_{h\pm}^{t1}$, there are three commonly used graph signal smoothness constraints.

1) The first one is $\mathrm{Tr}(\mathbf{Y}'\mathbf{L}_{h\pm}^{t1}\mathbf{Y}'^{\mathrm{T}})$, defined as

$$\mathrm{Tr}\big(\mathbf{Y}'\mathbf{L}_{h\pm}^{t1}\mathbf{Y}'^{\mathrm{T}}\big) = \frac{1}{2}\sum_{i,j=1}^{N_S} W_{h\pm,ij}^{t1}\big\|\mathbf{Y}_i' - \mathbf{Y}_j'\big\|_2^2 \quad (20)$$

where $\mathbf{L}_{h\pm}^{t1} = \mathbf{D}_{h\pm}^{t1} - (\mathbf{W}_{h\pm}^{t1} + (\mathbf{W}_{h\pm}^{t1})^{\mathrm{T}})/2$ denotes the conventional graph Laplacian matrix, with $\mathbf{D}_{h\pm}^{t1}$ being a diagonal degree matrix with $D_{h\pm,ii}^{t1} = (1/2)\sum_{j=1}^{N_S} W_{h\pm,ij}^{t1} + W_{h\pm,ji}^{t1}$. This smoothness prior promotes a small difference between the vertices connected by the positive edges with $W_{h\pm,ij}^{t1} > 0$, and a large difference between the vertices connected by the negative edges with $W_{h\pm,ij}^{t1} < 0$. Intuitively, this prior agrees with the structural consistency property. However, directly using this smoothness constraint will pose numerical problems in the transformation model. For example, it would result in $\mathbf{Y}_i' \to +\infty$ and $\mathbf{Y}_i' \to -\infty$ when $W_{h\pm,ij}^{t1}$ is negative.

2) The second commonly used constraint utilizes the signed graph Laplacian $\bar{L}_{h\pm}^{t1}$ defined in [36], where $\bar{L}_{h\pm}^{t1} = \bar{D}_{h\pm}^{t1} - (\mathbf{W}_{h\pm}^{t1} + (\mathbf{W}_{h\pm}^{t1})^{\mathrm{T}})/2$ with $\bar{D}_{h\pm}^{t1}$ being a diagonal degree matrix with $\bar{D}_{h\pm,ii}^{t1} = (1/2)\sum_{j=1}^{N_S} |W_{h\pm,ij}^{t1}| + |W_{h\pm,ji}^{t1}|$. Using this $\mathrm{Tr}(\mathbf{Y}'\bar{L}_{h\pm}^{t1}\mathbf{Y}'^{\mathrm{T}})$ as the smoothness constraint, we have

$$\mathrm{Tr}\big(\mathbf{Y}'\bar{L}_{h\pm}^{t1}\mathbf{Y}'^{\mathrm{T}}\big) = \frac{1}{2}\sum_{i,j=1}^{N_S} \big|W_{h\pm,ij}^{t1}\big|\big\|\mathbf{Y}_i' - \mathrm{sgn}\big(W_{h\pm,ij}^{t1}\big)\mathbf{Y}_j'\big\|_2^2$$

$$(21)$$

where $\mathrm{sgn}(\cdot)$ denotes the sign function. Since the signed Laplacian matrix $\bar{L}_{h\pm}^{t1}$ is positive-semidefinite, the function of (21) is convex, which is beneficial for the solution of the image transformation model. However, the penalty term of $\|\mathbf{Y}_i' - \mathrm{sgn}(W_{h\pm,ij}^{t1})\mathbf{Y}_j'\|_2^2$ with negative edge weight only requires that $\mathbf{Y}_i'$ and $\mathbf{Y}_j'$ be of the similar magnitude but of opposite signs. For example, $\mathbf{Y}_i' = \rho$ and $\mathbf{Y}_j' = -\rho$ will compute this penalty term to 0, which does not satisfy the requirement of structural consistency property.

3) The third alternative constraint uses a generalized graph Laplacian $\mathbf{L}_{h\pm,g}^{t1}$ defined in [43], where $\mathbf{L}_{h\pm,g}^{t1} = \mathbf{L}_{h\pm}^{t1} + \Xi$ and Ξ is a perturbation matrix to make $\mathbf{L}_{h\pm,g}^{t1}$ to be positive-semidefinite, and it is set to be $\Xi = \tau \mathbf{I}$ with $\tau = -\lambda_{\min}$ and $\lambda_{\min}$ being the smallest eigenvalue of $\mathbf{L}_{h\pm}^{t1}$. Using this $\mathrm{Tr}(\mathbf{Y}'\mathbf{L}_{h\pm,g}^{t1}\mathbf{Y}'^{\mathrm{T}})$ as the smoothness constraint, we have

$$\mathrm{Tr}\big(\mathbf{Y}'\mathbf{L}_{h\pm,g}^{t1}\mathbf{Y}'^{\mathrm{T}}\big) = \frac{1}{2}\sum_{i,j=1}^{N_S} W_{h\pm,ij}^{t1}\big\|\mathbf{Y}_i' - \mathbf{Y}_j'\big\|_2^2 + \tau\sum_{i=1}^{N_S}\big\|\mathbf{Y}_i'\big\|_2^2. \qquad (22)$$

It can be seen that to make the smoothness constraint numerically stable, (22) adds a feature energy penalty term to prevent $\mathbf{Y}_i' \to +\infty$ and $\mathbf{Y}_i' \to -\infty$ with negative $W_{h\pm,ij}^{t1}$ as in (20). At the same time, it ensures the convex property of the constraint function because $\mathbf{L}_{h\pm,g}^{t1}$ is positive-semidefinite. However, this smoothness constraint suffers from two drawbacks: first, it carries a bias for $\mathbf{Y}'$, i.e., it makes $\mathbf{Y}'$ tend to small amplitude values, which is physically unfounded and harmful to the image transformation model; second, the value of $\lambda_{\min}$ is difficult to determine, i.e., solving for the minimum eigenvalue of the matrix $\mathbf{L}_{h\pm}^{t1}$ introduces additional computational complexity.

APPENDIX B
OPTIMIZATION OF TRANSFORMATION MODEL

We utilize the ADMM to solve the transformation model of (17), whose augmented Lagrangian function is

$$\Theta\big(\mathbf{Y}', \Delta^y, \mathbf{R}\big)$$
$$= 2\mathrm{Tr}\big(\mathbf{Y}'\mathbf{L}_{h+}^{t1}\mathbf{Y}'^{\mathrm{T}}\big) + \alpha\sum_{i,j=1}^{N_S} W_{h-,ij}^{t1} f\Big(\big\|\mathbf{Y}_i' - \mathbf{Y}_j'\big\|_2\Big)$$
$$+ \beta\big\|\mathbf{Y}'\big\|_{\mathrm{RGL}} + \lambda\big\|\Delta^y\big\|_{2,1} + \mathrm{Tr}\big(\mathbf{R}^{\mathrm{T}}\big(\mathbf{Y}' - \mathbf{Y} - \Delta^y\big)\big)$$
$$+ \frac{\mu}{2}\big\|\mathbf{Y}' - \mathbf{Y} - \Delta^y\big\|_F^2 \qquad (23)$$

where $\mu > 0$ is a penalty parameter and $\mathbf{R} \in \mathbb{R}^{2B_2 \times N_S}$ is a Lagrange multiplier. The minimization problem of (23) can be written as three subproblems.

A. $\mathbf{Y}'$-Subproblem

Minimizing function (23) with respect to $\mathbf{Y}'$ is given by

$$\min_{\mathbf{Y}'} 2\mathrm{Tr}\big(\mathbf{Y}'\mathbf{L}_{h+}^{t1}\mathbf{Y}'^{\mathrm{T}}\big) + \alpha\sum_{i,j=1}^{N_S} W_{h-,ij}^{t1} f\Big(\big\|\mathbf{Y}_i' - \mathbf{Y}_j'\big\|_2\Big)$$
$$+ \beta\big\|\mathbf{Y}'\big\|_{\mathrm{RGL}} + +\frac{\mu}{2}\big\|\mathbf{Y}' - \mathbf{Y} - \Delta^y + \mathbf{R}/\mu\big\|_F^2. \qquad (24)$$

Using the gradient descent method, we can iteratively update $\mathbf{Y}'$ with an inner loop. Take penalty function $f(d) = -(1/(d^2 + \epsilon))$ as an example, we have

$$\varphi\big(\mathbf{Y}'\big) = 4\mathbf{Y}'\big(\mathbf{L}_{h+}^{t1} + \Psi\big) + \mu\big(\mathbf{Y}' - \mathbf{Y} - \Delta^y\big) + \mathbf{R}$$
$$\mathbf{Y}' \leftarrow \mathbf{Y}' - \xi\varphi\big(\mathbf{Y}'\big) \qquad (25)$$

where ξ is a step size, and $\Psi = \Psi^1 - \Psi^2$ with $\Psi^1 \in \mathbb{R}^{N_S \times N_S}$ being a diagonal matrix with $\Psi_{ii}^1 = \sum_{j=1}^{N_S}(\Psi_{ij}^2 + \Psi_{ji}^2)/2$ and Ψ_{ij}^2 being computed by

$$\Psi_{ij}^2 = \beta e^{-\big\|\mathbf{Y}_i' - \mathbf{Y}_j'\big\|_2^2/\sigma_{ij}^y} W_{h+,ij}^{t2}\Big(1 - \big\|\mathbf{Y}_i' - \mathbf{Y}_j'\big\|_2^2/\sigma_{ij}^y\Big)$$
$$- \alpha W_{h-,ij}^{t1}/\Big(\big\|\mathbf{Y}_i' - \mathbf{Y}_j'\big\|_2^2 + \epsilon\Big)^2. \qquad (26)$$

B. Δ^y-Subproblem

Minimizing function (23) with respect to Δ^y is given by

$$\min_{\Delta^y} \lambda\big\|\Delta^y\big\|_{2,1} + +\frac{\mu}{2}\big\|\Delta^y + \mathbf{Y} - \mathbf{Y}' - \mathbf{R}/\mu\big\|_F^2 \qquad (27)$$

whose closed-form solution is given by Yang et al. [44, Lemma 3.3] as

$$\Delta_i^y = \max\left\{\|\mathbf{Q}_i\|_2 - \frac{\lambda}{\mu}, 0\right\}\frac{\mathbf{Q}_i}{\|\mathbf{Q}_i\|_2} \qquad (28)$$

where $\mathbf{Q} = \mathbf{Y}' - \mathbf{Y} + \mathbf{R}/\mu$ and we follow the convention $0 \cdot (0/0) = 0$.

C. $\mathbf{R}$-Subproblem

Finally, the Lagrangian multiplier is updated by

$$\mathbf{R} \leftarrow \mathbf{R} + \mu\big(\mathbf{Y}' - \mathbf{Y} - \Delta^y\big). \qquad (29)$$

REFERENCES

[1] A. Singh, "Review article digital change detection techniques using remotely-sensed data," *Int. J. Remote Sens.*, vol. 10, no. 6, pp. 989–1003, Jun. 1989.

[2] Z. Lv, T. Liu, J. A. Benediktsson, and N. Falco, "Land cover change detection techniques: Very-high-resolution optical images: A review," *IEEE Geosci. Remote Sens. Mag.*, vol. 10, no. 1, pp. 44–63, Mar. 2022.

[3] L. Zhang and L. Zhang, "Artificial intelligence for remote sensing data analysis: A review of challenges and opportunities," *IEEE Geosci. Remote Sens. Mag.*, vol. 10, no. 2, pp. 270–294, Jun. 2022.

[4] Y. Sun, L. Lei, X. Li, X. Tan, and G. Kuang, "Structure consistency-based graph for unsupervised change detection with homogeneous and heterogeneous remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 4700221.

[5] R. Touati, "Détection de changement en imagerie satellitaire multimodale," Ph.D. dissertation, Département d'Informatique et de Recherche Opérationnelle, Université de Montréal, Montréal, QC, Canada, 2019.

[6] Z. Lv et al., "Land cover change detection with heterogeneous remote sensing images: Review, progress, and perspective," *Proc. IEEE*, vol. 110, no. 12, pp. 1976–1991, Dec. 2022.

[7] M. Volpi, G. Camps-Valls, and D. Tuia, "Spectral alignment of multi-temporal cross-sensor images with automated kernel canonical correlation analysis," *ISPRS J. Photogramm. Remote Sens.*, vol. 107, pp. 50–63, Sep. 2015.

[8] X. Niu, M. Gong, T. Zhan, and Y. Yang, "A conditional adversarial network for change detection in heterogeneous images," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 1, pp. 45–49, Jan. 2019.

[9] Y. Tang, X. Yang, T. Han, K. Sun, Y. Guo, and J. Hu, "Iterative optimization-enhanced contrastive learning for multimodal change detection," *Remote Sens.*, vol. 16, no. 19, p. 3624, Sep. 2024.

[10] L. Wan, Y. Xiang, and H. You, "An object-based hierarchical compound classification method for change detection in heterogeneous optical and SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 12, pp. 9941–9959, Dec. 2019.

[11] T. Han, Y. Tang, X. Yang, Z. Lin, B. Zou, and H. Feng, "Change detection for heterogeneous remote sensing images with improved training of hierarchical extreme learning machine (HELM)," *Remote Sens.*, vol. 13, no. 23, p. 4918, Dec. 2021.

[12] C. Li, G. Li, X. Wang, and P. K. Varshney, "A copula-based method for change detection with multisensor optical remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 5620015.

[13] L. Zhu, W. Sun, D. Fan, H. Xing, and X. Liu, "Unsupervised spatial self-similarity difference-based change detection method for multi-source heterogeneous images," *Pattern Recognit.*, vol. 149, May 2024, Art. no. 110237.

[14] J. Shi, T. Wu, A. Kai Qin, Y. Lei, and G. Jeon, "Self-guided autoencoders for unsupervised change detection in heterogeneous remote sensing images," *IEEE Trans. Artif. Intell.*, vol. 5, no. 6, pp. 2458–2471, Jun. 2024.

[15] Z. Lv, J. Liu, W. Sun, T. Lei, J. A. Benediktsson, and X. Jia, "Hierarchical attention feature fusion-based network for land cover change detection with homogeneous and heterogeneous remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 4411115.

[16] Z. Ma, R. Wang, F. Hao, and L. Song, "Heterogeneous image change detection based on dual image translation and dual contrastive learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 62, 2024, Art. no. 4704714.

[17] A. Dai, J. Yang, T. Zhang, B. Gao, K. Tang, and X. Chen, "DADR-HCD: A deep domain adaptation and disentangled representation network for unsupervised heterogeneous change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 62, 2024, Art. no. 5625315.

[18] M. Mignotte, "A fractal projection and Markovian segmentation-based approach for multimodal change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 11, pp. 8046–8058, Nov. 2020.

[19] L. T. Luppino, F. M. Bianchi, G. Moser, and S. N. Anfinsen, "Unsupervised image regression for heterogeneous change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 12, pp. 9960–9975, Dec. 2019.

[20] C. Li, G. Li, Z. Wang, X. Wang, and P. K. Varshney, "COMIC: An unsupervised change detection method for heterogeneous remote sensing images based on copula mixtures and cycle-consistent adversarial networks," *Inf. Fusion*, vol. 106, Jun. 2024, Art. no. 102240.

[21] L. T. Luppino et al., "Code-aligned autoencoders for unsupervised change detection in multimodal remote sensing images," *IEEE Trans. Neural Netw. Learn.*, vol. 35, no. 1, pp. 60–72, Jan. 2024.

[22] L. T. Luppino et al., "Deep image translation with an affinity-based change prior for unsupervised multimodal change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 4700422.

[23] H. Wu, J. Geng, and W. Jiang, "Multidomain constrained translation network for change detection in heterogeneous remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 62, 2024, Art. no. 5616916.

[24] Y. Sun, L. Lei, D. Guan, J. Wu, G. Kuang, and L. Liu, "Image regression with structure cycle consistency for heterogeneous change detection," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 2, pp. 1613–1627, Feb. 2024.

[25] H. Chen, N. Yokoya, and M. Chini, "Fourier domain structural relationship analysis for unsupervised multimodal change detection," *ISPRS J. Photogramm. Remote Sens.*, vol. 198, pp. 99–114, Apr. 2023.

[26] Y. Sun, L. Lei, D. Guan, and G. Kuang, "Iterative robust graph for unsupervised change detection of heterogeneous remote sensing images," *IEEE Trans. Image Process.*, vol. 30, pp. 6277–6291, 2021.

[27] T. Han, Y. Tang, B. Zou, and H. Feng, "Unsupervised multimodal change detection based on adaptive optimization of structured graph," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 126, Feb. 2024, Art. no. 103630.

[28] H. Chen, N. Yokoya, C. Wu, and B. Du, "Unsupervised multimodal change detection based on structural relationship graph representation learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5635318.

[29] Y. Tang, X. Yang, T. Han, F. Zhang, B. Zou, and H. Feng, "Enhanced graph structure representation for unsupervised heterogeneous change detection," *Remote Sens.*, vol. 16, no. 4, p. 721, Feb. 2024.

[30] Y. Sun, L. Lei, L. Liu, and G. Kuang, "Structural regression fusion for unsupervised multimodal change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 4504018.

[31] X. Zheng, D. Guan, B. Li, Z. Chen, and X. Li, "Change smoothness-based signal decomposition method for multimodal change detection," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022.

[32] J. F. Florez-Ospina, D. A. Jimenez-Sierra, H. D. Benitez-Restrepo, and G. R. Arce, "Exploiting variational inequalities for generalized change detection on graphs," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 4705616.

[33] Y. Sun, L. Lei, D. Guan, M. Li, and G. Kuang, "Sparse-constrained adaptive structure consistency-based unsupervised image regression for heterogeneous remote-sensing change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 4405814.

[34] Y. Sun, L. Lei, Z. Li, and G. Kuang, "Similarity and dissimilarity relationships based graphs for multimodal change detection," *ISPRS J. Photogramm. Remote Sens.*, vol. 208, pp. 70–88, Feb. 2024.

[35] D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega, and P. Vandergheynst, "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains," *IEEE Signal Process. Mag.*, vol. 30, no. 3, pp. 83–98, May 2013.

[36] J. Kunegis, S. Schmidt, A. Lommatzsch, J. Lerner, E. W. D. Luca, and S. Albayrak, "Spectral analysis of signed graphs for clustering, prediction and visualization," in *Proc. SIAM Int. Conf. Data Mining*, 2010, pp. 559–570.

[37] L. Chu, Z. Wang, J. Pei, J. Wang, Z. Zhao, and E. Chen, "Finding gangs in war from signed networks," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2016, pp. 1505–1514.

[38] Z. Ban, J. Liu, and L. Cao, "Superpixel segmentation using Gaussian mixture model," *IEEE Trans. Image Process.*, vol. 27, no. 8, pp. 4105–4117, Aug. 2018.

[39] T. Han, Y. Tang, Y. Chen, X. Yang, Y. Guo, and S. Jiang, "SDC-GAE: Structural difference compensation graph autoencoder for unsupervised multimodal change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 62, 2024, Art. no. 5622416.

[40] F. Heider, "Attitudes and cognitive organization," *J. Psychol.*, vol. 21, no. 1, pp. 107–112, Jan. 1946.

[41] D. Cartwright and F. Harary, "Structural balance: A generalization of heider's theory," *Psychol. Rev.*, vol. 63, no. 5, p. 277, Aug. 1956.

[42] D. A. Jimenez-Sierra, H. D. Benítez-Restrepo, H. D. Vargas-Cardona, and J. Chanussot, "Graph-based data fusion applied to: Change detection and biomass estimation in Rice crops," *Remote Sens.*, vol. 12, no. 17, p. 2683, Aug. 2020.

[43] G. Cheung, W.-T. Su, Y. Mao, and C.-W. Lin, "Robust semisupervised graph classifier learning with negative edge weights," *IEEE Trans. Signal Inf. Process. Over Netw.*, vol. 4, no. 4, pp. 712–726, Dec. 2018.

[44] J. Yang, W. Yin, Y. Zhang, and Y. Wang, "A fast algorithm for edge-preserving variational multichannel image restoration," *SIAM J. Imag. Sci.*, vol. 2, no. 2, pp. 569–592, Jan. 2009.

Yuli Sun received the Ph.D. degree in information and communication engineering from the National University of Defense Technology (NUDT), Changsha, China, in 2023.

He is currently a Post-Doctoral Researcher with the College of Aerospace Science and Engineering, NUDT. He has published papers in IEEE TRANSACTIONS ON IMAGE PROCESSING (TIP), IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS (TNNLS), IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING (TGRS), ISPRS *Journal of Photogrammetry and Remote Sensing* (P&RS), and *Pattern Recognition*. His research interests cover remote sensing image processing and multitemporal image analysis.

Ming Li received the M.S. degree from Central South University, Changsha, China, in 2017, and the Ph.D. degree in information and communication engineering from the National University of Defense Technology (NUDT), Changsha, in 2022.

He is currently a Lecturer with the Information and Communication College, NUDT. In addition, he is a Post-Doctoral Researcher with the College of Electronic Science, NUDT. His research interests include remote sensing scene interpretation, multimodal data fusion, and object detection.

Zhang Li received the Ph.D. degree from Delft University of Technology, Delft, The Netherlands, in 2015.

He is an Associate Professor with the College of Aerospace Science and Engineering, National University of Defense Technology, Changsha, China. He has authored more than 30 papers in high-ranking journals and conferences. His research interests include biomedical image analysis, particularly image registration, and vision navigation.

Lin Lei received the Ph.D. degree in information and communication engineering from the National University of Defense Technology, Changsha, China, in 2008.

She is currently a Professor with the School of Electronic Science, National University of Defense Technology. Her research interests include computer vision, remote sensing image interpretation, and data fusion.

Gangyao Kuang (Senior Member, IEEE) received the B.S. and M.S. degrees in geophysics from the Central South University of Technology, Changsha, China, in 1988 and 1991, respectively, and the Ph.D. degree in communication and information from the National University of Defense Technology, Changsha, in 1995.

He is currently a Professor at the School of Electronic Science, National University of Defense Technology. His research interests include remote sensing, SAR image processing, change detection, SAR ground moving target indication, and classification with polarimetric SAR images.