

Image Regression With Structure Cycle Consistency for Heterogeneous Change Detection

Yuli Sun^{ID}, Lin Lei^{ID}, Dongdong Guan, Junzheng Wu, Gangyao Kuang, and Li Liu^{ID}, *Senior Member, IEEE*

Abstract—Change detection (CD) between heterogeneous images is an increasingly interesting topic in remote sensing. The different imaging mechanisms lead to the failure of homogeneous CD methods on heterogeneous images. To address this challenge, we propose a structure cycle consistency-based image regression method, which consists of two components: the exploration of structure representation and the structure-based regression. We first construct a similarity relationship-based graph to capture the structure information of image; here, a k -selection strategy and an adaptive-weighted distance metric are employed to connect each node with its truly similar neighbors. Then, we conduct the structure-based regression with this adaptively learned graph. More specifically, we transform one image to the domain of the other image via the structure cycle consistency, which yields three types of constraints: forward transformation term, cycle transformation term, and sparse regularization term. Noteworthy, it is not a traditional pixel value-based image regression, but an image structure regression, i.e., it requires the transformed image to have the same structure as the original image. Finally, change extraction can be achieved accurately by directly comparing the transformed and original images. Experiments conducted on different real datasets show the excellent performance of the proposed method. The source code of the proposed method will be made available at <https://github.com/yulisun/AGSCC>.

Index Terms—Adaptive graph (AG), cycle consistency, heterogeneous, image regression, multimodal, unsupervised change detection (CD).

NOMENCLATURE

$\tilde{\mathbf{X}}, \tilde{\mathbf{Y}}$	Preevent and postevent images.
$\tilde{\mathbf{X}}', \tilde{\mathbf{Y}}'$	Regression images of $\tilde{\mathbf{Y}}$ and $\tilde{\mathbf{X}}$.
$\tilde{\mathbf{x}}_i$	i th superpixel of $\tilde{\mathbf{X}}$.
\mathbf{X}, \mathbf{Y}	Feature matrices of $\tilde{\mathbf{X}}$ and $\tilde{\mathbf{Y}}$.
\mathbf{X}', \mathbf{Y}'	Regression feature matrices of \mathbf{Y} and \mathbf{X} .
$\tilde{\mathbf{X}}''$	Back transformed image of $\tilde{\mathbf{Y}}'$.

Manuscript received 26 September 2021; revised 11 March 2022 and 7 May 2022; accepted 12 June 2022. Date of publication 29 June 2022; date of current version 6 February 2024. This work was supported in part by the National Key Research and Development Program of China under Grant 2021YFB3100800; in part by the National Natural Science Foundation of China under Grant 61872379, Grant 12171481, and Grant 61971426; and in part by the Natural Science Basic Research Plan in Shaanxi Province of China under Grant 2022JQ-694. (*Corresponding author: Lin Lei.*)

Yuli Sun, Lin Lei, Junzheng Wu, and Gangyao Kuang are with the College of Electronic Science, National University of Defense Technology, Changsha 410073, China (e-mail: alaleilin@163.com).

Dongdong Guan is with the High-Tech Institute of Xi'an, Xi'an 710025, China.

Li Liu is with the College of System Engineering, National University of Defense Technology, Changsha 410073, China.

This article has supplementary downloadable material available at <https://doi.org/10.1109/TNNLS.2022.3184414>.

Digital Object Identifier 10.1109/TNNLS.2022.3184414

\mathbf{X}''	Back transformed feature matrix of \mathbf{Y}' .
$\mathbf{X}^{(m)}$	m th feature matrix of $\tilde{\mathbf{X}}$.
$\mathbf{X}_i^{(m)}$	m th feature vector of the i th superpixel $\tilde{\mathbf{x}}_i$.
$dist_{m,i,j}^X$	m th feature distance between $\tilde{\mathbf{x}}_i$ and $\tilde{\mathbf{x}}_j$.
\mathbf{S}^X	Weighting matrix representing the structure of $\tilde{\mathbf{X}}$.
w_m^X	Weight of the m th feature of $\mathbf{X}^{(m)}$.
$\mathbf{L}_X, \mathbf{L}_Y$	Graph Laplacian matrices.
Δ	Changed feature matrix.
\mathbf{I}_N	$N \times N$ identity matrix.
$\mathbf{1}_N$	$N \times 1$ column vector of ones.

I. INTRODUCTION

A. Background

CHANGE detection (CD) is a well-known task in remote sensing (RS), which is to identify the changes on Earth's surface by comparing images acquired over the same area but at different times [1]. CD has been widely used in many applications, including environmental monitoring, agricultural surveys, and disaster assessment [2]–[4].

Currently, most CD algorithms are based on the homogeneous images, that is, the multitemporal images are acquired from the same sensor, such as homogeneous CD of optical images [5], synthetic aperture radar (SAR) images [6], [7], and hyperspectral images [8], [9]. However, in some cases, homogeneous images cannot be acquired in time, and we have to focus on heterogeneous CD, i.e., using images from different sensors to detect changes, which roughly contains two types: cross-sensor or multisensor images acquired by different sensors with the same sensor type (such as two multispectral images from Sentinel-2 and Landsat-8) and multisource images acquired by different types of sensors (such as a pair of optical and SAR images). Heterogeneous CD is particularly attractive in two aspects. First, it can improve the temporal resolution of change trend analysis or extend the time frame of long-term monitoring by inserting heterogeneous data along the timeline [10]. Second, it can shorten the response time of CD analysis in emergency natural disasters (such as floods and earthquakes) and rescue operations. In such scenarios, heterogeneous CD allows to use the first acquired image to analyze changes without having to wait for homogeneous images, and more importantly, sometimes the acquired homogeneous images may be unavailable due to the accompanying adverse light and weather conditions.

Since the heterogeneous images usually have different representations for the same object and show different statistical properties, they can not be directly compared to generate the

difference image (DI) as in homogeneous CD. Therefore, the core of heterogeneous CD is to make the “incomparable” images “comparable.” Let \mathbf{X} and \mathbf{Y} be the coregistered images to be compared and x and y be the data samples at the same position in \mathbf{X} and \mathbf{Y} , respectively, which can be pixels (i.e., the smallest item of information in an image), square patches (i.e., a small square piece of an image), or superpixels (i.e., a group of pixels that are similar in color and other low-level properties), according to the CD method’s basic unit of analysis. In the coregistered images, x and y represent the same position on the ground. In the homogeneous CD, we can directly compare x and y to measure the change level by using some algebraic operations, such as image differencing [11] and image ratioing/log ratioing [12]. In the heterogeneous CD, we need to transform \mathbf{X} and \mathbf{Y} into the same domain to make x and y comparable, which relates to the topic of image transformation.

Depending on whether training samples are required in the transformation process, heterogeneous CD methods can be classified as supervised or unsupervised; according to the method used for transformation, heterogeneous CD methods can be divided into deep learning-based and traditional signal processing-based. Meanwhile, based on the transformed common domain, heterogeneous CD methods can also be roughly divided into three categories as follows.

- 1) *Image Classification-Based Methods:* In these methods, the heterogeneous images are transformed into a common category space. They detect the changes by comparing the classification results of multitemporal images, such as the postclassification comparison (PCC) [13], multitemporal segmentation and compound classification (MS-CC) [14], and the classified adversarial network-based method (CAN) [15]. The advantage of such methods is that they are intuitive and easy to implement, and they can provide categories of changes (indicating what kind of change has occurred) along with the detection of changes. However, they also have the following drawbacks: the CD accuracy is limited by the accuracy and precision of classification, and there is also the risk of suffering from the accumulation of classification errors.
- 2) *Feature Transformation-Based Methods:* In these methods, the heterogeneous images are transformed into a common constructed or learned feature space. Some of them manually construct the modality-invariant similarity measures, such as Kullback–Leibler (KL) distance based on copula theory [16], manifold distance [17], kernel canonical correlation analysis (kCCA) [18], and non-local pixel pairwise energy-based model (NLPEM) [19]. Some of them learn the latent features of heterogeneous images by using deep neural networks (DNNs), such as symmetric convolutional coupling network (SCCN) [20], probabilistic model based on bipartite convolutional neural network (BCCN) [21], spatially self-paced convolutional network (SSPCN) [22], commonality autoencoder-based common feature learning (CACFL) [23], and transfer learning-based semisupervised Siamese network (S3N) [24]. For these feature

transformation-based methods, the CD performance relies on the extracted features, and they often face these challenges: it is difficult to obtain the common feature when the scene is very complex (the ground features vary greatly), the noise in image is severe (especially for SAR image), or the training samples are not enough or mixed with wrong samples.

- 3) *Image Regression-Based Methods:* These methods transform one image to the domain of the other image, which can be regarded as the image-to-image translation. The homogeneous pixel transformation (HPT) method [25] builds up the relationship between pixel values of heterogeneous images by using a k -nearest neighbor (KNN)-based multivalue estimation strategy, which is supervised by the unchanged pixels. To avoid the reliance on unchanged samples, an affinity matrix difference-based image regression (AMD-IR) method is proposed [26], which uses AMD to pick pixels that have a high probability of being unchanged as the pseudo-training set. Some DNN-based regression methods have also been proposed, such as the deep translation-based CD network (DTCDN) [27], conditional generative adversarial network (cGAN) [28], AMD-based X-Net and adversarial cyclic encoder network (ACE-Net) [29], cycle-consistent adversarial network (CycleGAN)-based [30] unsupervised change detection network (USCDN) [31], and the image style transfer (IST)-based method [32]. Since the regression process usually needs to be trained with unchanged pairs of heterogeneous data (pixels or square patches), these regression-based methods either require a labeled training set (such as HPT) or the pseudo-training set/change prior to guide the regression process (such as AMD-IR, ACE-Net, and X-Net) or use a complex iterative coarse-to-fine process to filter the training data (such as cGAN and IST).

Generally, since reference samples are often not available in practical applications, the unsupervised CD methods seem to be more attractive than supervised methods [22], [33]. At the same time, the datasets for heterogeneous CD are relatively limited until now, which is because constructing a ground-truth map that reflects real change information requires a high cost of manual operation and great expert knowledge in practice. There is currently no large publicly available dataset to support a pretrained model for heterogeneous CD, and most deep learning-based methods detect the changes based on the preevent and postevent images themselves [20], [21], [23], [27], [29]. Therefore, the traditional unsupervised heterogeneous CD methods are still very appealing: on the one hand, they can quickly and automatically extract change information; and on the other hand, they can provide assistance to deep learning-based methods, such as supporting the training process [29] or constructing high confidence pseudo-training sets, similar to the unsupervised CD methods proposed in [8] and [34].

B. Motivations

In this article, we aim to propose an unsupervised image regression-based method for heterogeneous CD, which is

based on the structure consistency between heterogeneous images. Recently, the structure features have been used as modality-invariant similarity metrics by transformation-based methods [35], [36], which compare the structures of heterogeneous images in the same domain to calculate the DI by graph projection. Meanwhile, instead of directly focusing on the structure difference between images, some regression-based methods first use the structure consistency to complete the image translation and then compare the translated preevent (or postevent) image and original postevent (or preevent) image to calculate the DI, such as the fractal projection and Markovian segmentation (FPMS)-based method [37], patch similarity graph matrix-based method (PSGM) [38], and sparse constrained adaptive structure consistency (SCASC)-based method [39].

FPMS [37], PSGM [38], and SCASC [39] translate the preevent image to the domain of postevent image as $\mathbf{X} \xrightarrow{\mathbf{S}} \mathbf{Y}'$ by fractal projection with fractal code \mathbf{S} , image reconstruction with self-expression graph matrix \mathbf{S} , and image decomposition with adaptive probabilistic graph \mathbf{S} , respectively, where these \mathbf{S} can represent the relationships between image patches/blocks. Specifically, structural consistency in FPMS [37] is expressed as follows. If image \mathbf{X} can be encoded by the fractal code \mathbf{S} , then the regression image can be decoded by fractal projection with image \mathbf{Y} and \mathbf{S} . In PSGM [38], the structural consistency is expressed as follows. If the preevent image can be represented by $\mathbf{X} = \mathbf{XS}$ with the self-expression graph matrix \mathbf{S} (here, \mathbf{X} is the patch group matrix), then the unchanged postevent image should conform to this structure \mathbf{S} and also satisfy $\mathbf{Y} = \mathbf{YS}$. Meanwhile, the structural consistency in SCASC [39] can be expressed by the fact that if \mathbf{X}_i and \mathbf{X}_j in the preevent image are very similar, \mathbf{Y}_i and \mathbf{Y}_j in the postevent image should also be similar in the absence of changes (here, \mathbf{X}_i and \mathbf{Y}_i and \mathbf{X}_j and \mathbf{Y}_j represent superpixels at the same position, respectively), which means that the similarity relationship between superpixels within the image is consistent across different modalities. Recently, a structured graph-based image regression with Markovian segmentation (GIR-MRF) [40] has been proposed, which can be regarded as a combination of PSGM and SCASC to capture both global and local structural information, thereby improving image regression and CD performance.

In these heterogeneous CD methods with structure-based image regression, there are two key aspects: first, how to construct \mathbf{S} to capture the image structure, and second, how to complete the regression to detect the changes. Although these existing structure-based image regression methods achieve relatively good results [37]–[40], they still have two limitations as follows.

First, the constructed graph is not robust enough and it is difficult to fully characterize the structure of image. For example, the self-expression-based graph \mathbf{S} in PSGM [38] uses the global ℓ_1 -norm $\|\mathbf{S}\|_1$ to guarantee the sparsity of \mathbf{S} , but it ignores the local properties of each node. The SCASC [39] and GIR-MRF [40] use the KNN probabilistic graph to connect each node with its KNN to ensure sparsity and capture the image structure, but they directly use the Euclidean distance

metric in the graph construction and ignore the differences between different features. In this article, we propose an adaptive graph (AG) to capture the structure information of image by treating each superpixel as a node, which belongs to a KNN graph. To address the two challenges of the KNN graph, the choice of k and the distance metric, first, we use an in-degree-based adaptive k -selection strategy to choose a suitable k for each node instead of a fixed k in the common KNN graph; second, we use an adaptive-weighted distance metric that automatically determines the feature weights in the distance between superpixels, instead of a common Euclidean distance metric that gives the same weight to all the features without considering the distinguishability of different features in different images. Therefore, by combining the adaptive k and adaptive-weighted distance metric, we can find the truly similar neighbors for each node and assign suitable connection weights, thus better characterizing the image structure.

Second, the previous structure-based image regression methods of PSGM [38], SCASC [39], GIR-MRF [40], and FPMS [37] only consider one-way transformation of $\mathbf{X} \rightarrow \mathbf{Y}'$ or $\mathbf{Y} \rightarrow \mathbf{X}'$ and do not consider the inverse transformations of $\mathbf{Y}' \rightarrow \mathbf{X}$ or $\mathbf{X}' \rightarrow \mathbf{Y}$, where \mathbf{X}' and \mathbf{Y}' denote the translated images. Inspired by the cycle GAN [30], which can perform a compound translation of data from domain \mathcal{X} to domain \mathcal{Y} and then to domain \mathcal{X} (i.e., $\mathbf{X} \rightarrow \mathbf{Y}' \rightarrow \mathbf{X}$), we propose a structure cycle consistency-based image regression model. We decompose the postevent image into a translated image and a changed image by using three types of regularization terms: a forward transformation term ($\mathbf{X} \rightarrow \mathbf{Y}'$) to constrain the preevent image and translated image to have the same structure; a cycle transformation term to enable \mathbf{Y}' to be retransformed back to the original domain of \mathbf{X} as $\mathbf{Y}' \rightarrow \mathbf{X}$; and a sparse regularization term based on the prior sparse knowledge of changes that only a small part of the area changed and most areas remain unchanged during the event in the CD problem. By combining these regularization terms, we can obtain a better regression image and a more accurate DI with less noise, which further improves the CD performance.

C. Contributions

The main contributions of our work can be summarized as follows.

- 1) We propose an image regression-based heterogeneous CD method by using adaptive graph and structure cycle consistency (AGSCC), which combines two components: the exploration of structure representation and the structure-based regression.
- 2) We construct an adaptive distance-induced probabilistic graph to better capture the structure information of image, which combines the adaptive k -selection strategy and adaptive-weighted distance metric.
- 3) We design a structure cycle consistency-based image regression model by combining three types of regularization terms: forward transformation term, cycle transformation term, and sparse regularization term.

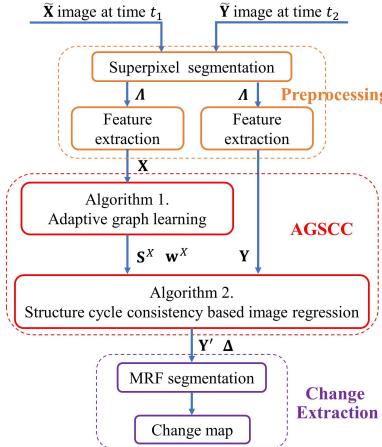


Fig. 1. Framework of the proposed heterogeneous CD method.

- 4) Comprehensive experiments on six real datasets verify the superiority of the proposed method compared with some state-of-the-art (SOTA) methods.

II. STRUCTURE CONSISTENCY

Assume that two heterogeneous images of $\tilde{\mathbf{X}} \in \mathbb{R}^{I \times J \times C_X}$ and $\tilde{\mathbf{Y}} \in \mathbb{R}^{I \times J \times C_Y}$ are acquired by different sensors over the same region before and after a change event, respectively. Here, I , J , and C_X (C_Y) represent the height, width, and number of channels of image $\tilde{\mathbf{X}}$ ($\tilde{\mathbf{Y}}$), respectively. We denote the pixel of each image as $\tilde{x}(i, j, c)$ and $\tilde{y}(i, j, c)$. The purpose of CD is to find the changed regions from the heterogeneous images. However, since $\tilde{\mathbf{X}}$ and $\tilde{\mathbf{Y}}$ are obtained by different sensors, they have different representations for the same object and show quite different statistical properties. Therefore, directly comparing their pixel values is meaningless.

In this article, we use the structure consistency to establish the connection between heterogeneous images. We first segment each image into small parts with the same segmentation map and use the similarity relationships between the image blocks within the same image to represent the image structure. Then, the structure consistency can be expressed as follows. If $\tilde{\mathbf{X}}_i$ and $\tilde{\mathbf{X}}_j$ represent the same kind of object (showing that they are very similar) and neither of them changes during the event, then $\tilde{\mathbf{Y}}_i$ and $\tilde{\mathbf{Y}}_j$ also represent the same kind of object (also showing that they are very similar). Because this nonlocal similarity within the image itself could eliminate the discrepancy between heterogeneous images, the similarity relationship-based structure can be well preserved across different imaging modalities. Therefore, we can use this structure consistency between heterogeneous images to translate the preevent image to the domain of postevent image, then calculate the DI to measure the change level, and finally compute the change map (CM).

The proposed method consists of four steps: 1) preprocessing, which includes superpixel segmentation and feature extraction; 2) structure representation by learning graph; 3) image translation with structure consistency; and 4) binary CM computation by segmentation or clustering methods, as shown in Fig. 1.

III. ADAPTIVE GRAPH

A. Preprocessing

As aforementioned, we need to consider the pairwise similarity relationships to represent image structures. Therefore, the superpixel that internally represents the same object is chosen as the basic analysis unit in AGSCC, which can bring two benefits compared to the individual pixel or square Patch. First, superpixel can maintain the shape and edge of object and contain contextual information. Second, it can greatly reduce the computational complexity by reducing the size of graph.

The simple linear iterative clustering (SLIC) method [41] is selected to generate the superpixels for its superiority in terms of efficiency (linear complexity of the number of image pixels) and boundary protection. SLIC is easy to use, and it offers flexibility in terms of the compactness and number of generated superpixels. For different types of preevent image $\tilde{\mathbf{X}}$, such as optical image with RGB bands, multispectral image with $C_X > 3$, and SAR image, we have some adjustments on the SLIC to generate the superpixels as in SCASC [39]. Once the preevent image $\tilde{\mathbf{X}}$ is segmented into N superpixels as $\tilde{\mathbf{X}}_n$, $n = 1, \dots, N$, we can obtain the segmentation map $\Lambda = \{\Lambda_n | n = 1, \dots, N\}$ with $\Lambda_n = \{(i, j) | \tilde{x}(i, j, c) \in \tilde{\mathbf{X}}_n, c = 1, \dots, C_X\}$. Then, Λ is mapped to the postevent image $\tilde{\mathbf{Y}}$ to generate the superpixels $\tilde{\mathbf{Y}}_n$, $n = 1, \dots, N$ with $\tilde{\mathbf{Y}}_n = \{\tilde{y}(i, j, c) | (i, j) \in \Lambda_n, c = 1, \dots, C_Y\}$, which can make $\tilde{\mathbf{X}}_n$ and $\tilde{\mathbf{Y}}_n$ represent the same region.

After the superpixel segmentation, different kinds of feature information can be extracted to represent the superpixel, such as the spatial, spectral (intensity), and textural information. We extract M features for each band and denote the m th feature extraction operator as $\mathcal{F}^{(m)}$, and then, we can obtain the feature vectors of $\mathbf{X}_n^{(m)} = \mathcal{F}^{(m)}(\tilde{\mathbf{X}}_n) \in \mathbb{R}^{C_X}$ and $\mathbf{Y}_n^{(m)} = \mathcal{F}^{(m)}(\tilde{\mathbf{Y}}_n) \in \mathbb{R}^{C_Y}$. By stacking these feature vectors, we can obtain the stacked feature matrices of $\mathbf{X} \in \mathbb{R}^{M \times C_X \times N}$ and $\mathbf{Y} \in \mathbb{R}^{M \times C_Y \times N}$, where the m th feature matrices are denoted as $\mathbf{X}^{(m)} = \mathbf{X}_{m,:,:} \in \mathbb{R}^{C_X \times N}$ and $\mathbf{Y}^{(m)} = \mathbf{Y}_{m,:,:} \in \mathbb{R}^{C_Y \times N}$.

B. Structure Representation by AG

1) *Graph Learning Model:* In order to characterize the image structure that is represented by the similarity relationship between superpixels, we first construct a weighted directed graph $\mathbf{G} = (\mathbf{V}, \mathbf{E}, \mathbf{S}^X)$ by setting each superpixel as a vertex, where the vertex set is $\mathbf{V} = \{\tilde{\mathbf{X}}_i | i = 1, \dots, N\}$ and the edge set is $\mathbf{E} = \{(\tilde{\mathbf{X}}_i, \tilde{\mathbf{X}}_j) | S_{i,j}^X \neq 0; i, j = 1, \dots, N\}$ with $S_{i,j}^X$ being the weight from $\tilde{\mathbf{X}}_i$ to $\tilde{\mathbf{X}}_j$. It is intuitive that the weight $S_{i,j}^X$ is larger if the distance between $\tilde{\mathbf{X}}_i$ and $\tilde{\mathbf{X}}_j$ is smaller. In [42], a distance-induced probabilistic graph is used, which learns the probability matrix \mathbf{S}^X by solving the following minimization model:

$$\min_{\mathbf{S}^X} \sum_{i,j=1}^N \text{dist}_{i,j}^X S_{i,j}^X + \alpha (S_{i,j}^X)^2, \quad \text{s.t. } \sum_{i=1}^N S_{i,j}^X = 1, \quad S_{i,j}^X \geq 0 \quad (1)$$

where $\alpha > 0$ is a tuning parameter and $\text{dist}_{i,j}^X$ is the distance between $\tilde{\mathbf{X}}_i$ and $\tilde{\mathbf{X}}_j$. As will be shown latter, \mathbf{S}^X can be

column k -sparse when the appropriate α is chosen. From (1), there are two points that directly affect the performance of this KNN-type graph: the α that controls the number (k) of NNs and the distance metric $dist^X$ that affects the weight and neighbor location of \mathbf{S}^X . In other words, α determines how many nonzero elements there are in \mathbf{S}^X , and $dist^X$ determines the position and values of each nonzero element in \mathbf{S}^X .

The Euclidean distance is often used directly for constructing the graph [39], [43], [44]. However, since different features represent different information, it is not appropriate to directly use the Euclidean distance metric that gives the same weight to all the features without considering the distinguishability of different features in different images. On the other hand, since the proportions of different kinds of objects in the image are also different, a fixed α for all superpixels that generates a common number k of NNs is also inappropriate, which may underconnect those superpixels that represent dominant occupied objects and overconnect those superpixels that represent nondominant occupied objects.

To address these challenges, we propose an AG that selects the appropriate neighbors for each node with adaptive k and adaptive-weighted distance metric by using the following minimization model:

$$\begin{aligned} & \min_{\mathbf{S}^X, \mathbf{w}^X} \sum_{i,j=1}^N \left(\sum_{m=1}^M w_m^X dist_{m,i,j}^X S_{i,j}^X \right) + \alpha_i (S_{i,j}^X)^2, \\ & \text{s.t. } \sum_{i=1}^N S_{i,j}^X = 1, \quad S_{i,j}^X \geq 0, \quad \sum_{m=1}^M (w_m^X)^\eta = 1, \quad w_m^X \geq 0 \end{aligned} \quad (2)$$

where $dist_{m,i,j}^X = \|\mathbf{X}_i^{(m)} - \mathbf{X}_j^{(m)}\|_2^2$ is the m th feature distance between $\tilde{\mathbf{X}}_i$ and $\tilde{\mathbf{X}}_j$, w_m^X is the weight for the m th feature distance, and $\alpha_i > 0$ and $0 < \eta < 1$ are tuning parameters. Compared to model (1), the graph learning model (2) introduces adaptive parameters α_i for each node and w_m^X for each feature, which are used to select k and the distance metric.

Define $\mathbf{1}_N$ as the $N \times 1$ column vector of ones, \mathbf{S}_i^X as the i th column of \mathbf{S}^X that represents the local structure of the i th superpixel, and $dist_{m,i}^X$ as the feature distance vector with the j th element being $dist_{m,i,j}^X$; problem (2) can be rewritten as the following model:

$$\begin{aligned} & \min_{\mathbf{S}^X, \mathbf{w}^X} \sum_{i=1}^N (\mathbf{S}_i^X)^T \left(\sum_{m=1}^M w_m^X dist_{m,i}^X \right) + \alpha_i \|\mathbf{S}_i^X\|_2^2 \\ & \text{s.t. } \mathbf{1}_N^T \mathbf{S}_i^X = 1, \quad S_{i,j}^X \geq 0, \quad \sum_{m=1}^M (w_m^X)^\eta = 1, \quad w_m^X \geq 0. \end{aligned} \quad (3)$$

The smooth term of $\alpha_i \|\mathbf{S}_i^X\|_2^2$ is used to avoid trivial solution and make \mathbf{S}_i^X sparse together with the conditions of $\mathbf{1}_N^T \mathbf{S}_i^X = 1$, $S_{i,j}^X \geq 0$. Specially, if we set $\alpha_i = 0$, problem (3) has a trivial solution for \mathbf{S}_i^X of $S_{i,i}^X = 1$ and $S_{j,i}^X = 0$, $j \neq i$, that is, $\tilde{\mathbf{X}}_i$ is only connected with itself with probability 1. On the other hand, if we set $\alpha_i \rightarrow \infty$, the optimal solution for \mathbf{S}_i^X is that $S_i^X = \mathbf{1}_N/N$, that is, all the superpixels connect $\tilde{\mathbf{X}}_i$ with the same probability $1/N$. The distance-induced regularization term enables neighbors with small distances to gain larger weights, the weighting vector \mathbf{w}^X is used to distinguish the

contribution of different features in the graph construction, and the parameter η is used to control the distribution of weights.

2) *Optimization*: Problem (3) can be efficiently solved by using the alternating direction method of multipliers (ADMM), which can iteratively update one variable at a time and fix the others. ADMM separates (3) into \mathbf{S}^X -subproblem and \mathbf{w}^X -subproblem.

a) \mathbf{S}^X -subproblem: Since each column of \mathbf{S}^X is independent with fixed \mathbf{w}^X in model (3), by defining $\mathbf{d}_i^X = (\sum_{m=1}^M w_m^X dist_{m,i}^X)$, the minimization of (3) with respect to \mathbf{S}^X can be reformulated as

$$\min_{\mathbf{1}_N^T \mathbf{S}_i^X = 1, S_{i,j}^X \geq 0} \left\| \mathbf{S}_i^X + \frac{1}{2\alpha_i} \mathbf{d}_i^X \right\|_2^2. \quad (4)$$

The Lagrangian function of (4) is

$$\mathcal{L}(\mathbf{S}_i^X, \zeta, \boldsymbol{\varsigma}) = \left\| \mathbf{S}_i^X + \frac{1}{2\alpha_i} \mathbf{d}_i^X \right\|_2^2 - 2\zeta (\mathbf{1}_N^T \mathbf{S}_i^X - 1) - \boldsymbol{\varsigma}^T \mathbf{S}_i^X \quad (5)$$

where $\zeta \in \mathbb{R}$ and $\boldsymbol{\varsigma} \in \mathbb{R}^N$ are two Lagrangian multipliers. With the Karush–Kuhn–Tucker (KKT) condition, the closed-form solution of \mathbf{S}_i^X is given by

$$\mathbf{S}_i^X = \left(-\frac{1}{2\alpha_i} \mathbf{d}_i^X + \zeta \right)_+ \quad (6)$$

where the operator $(\mathbf{Z})_+$ turns negative elements in \mathbf{Z} to zero while keeping the rest unchanged. We can find that \mathbf{S}_i^X is a sparse vector, which means that each superpixel is only connected with superpixels that are similar to itself. Then, we assume that \mathbf{S}_i^X has k_i nonzero elements. We sort \mathbf{d}_i^X in ascending order denoted as $d_{(1),i}^X, d_{(2),i}^X, \dots, d_{(N),i}^X$. By using (6), the following inequalities hold:

$$\begin{cases} -\frac{d_{(k_i),i}^X}{2\alpha_i} + \zeta > 0 \\ -\frac{d_{(k_i+1),i}^X}{2\alpha_i} + \zeta \leq 0. \end{cases} \quad (7)$$

Substituting the constraint of $\mathbf{1}_N^T \mathbf{S}_i^X = 1$ into (6), we have

$$\sum_{h=1}^{k_i} \left(-\frac{d_{(h),i}^X}{2\alpha_i} + \zeta \right) = 1. \quad (8)$$

Substituting (8) into (7), we have

$$\begin{cases} \alpha_i > \frac{k_i}{2} d_{(k_i),i}^X - \frac{1}{2} \sum_{h=1}^{k_i} d_{(h),i}^X \\ \alpha_i \leq \frac{k_i}{2} d_{(k_i+1),i}^X - \frac{1}{2} \sum_{h=1}^{k_i} d_{(h),i}^X. \end{cases} \quad (9)$$

Then, we can find that the regularization parameter α_i can be replaced by the number of neighbors k_i when we set

$$\alpha_i = \frac{k_i}{2} d_{(k_i+1),i}^X - \frac{1}{2} \sum_{h=1}^{k_i} d_{(h),i}^X. \quad (10)$$

Therefore, the value of α_i can be determined by k_i , that is, the tuning of parameter α_i is replaced by the tuning of k_i , which is more intuitive (k_i has explicit meaning) and

easier (k_i is an integer). With the number k_i of NNs, the \mathbf{S}^X -subproblem can be solved as

$$S_{(j),i}^X = \begin{cases} \frac{d_{(k_i+1),i}^X - d_{(j),i}^X}{k_i d_{(k_i+1),i}^X - \sum_{h=1}^{k_i} d_{(h),i}^X}, & j \leq k_i \\ 0, & j > k_i. \end{cases} \quad (11)$$

It can be found that \mathbf{S}_i^X is a sparse vector with k_i nonzero elements, which means that $\tilde{\mathbf{X}}_i$ is connected to k_i NNs with the weight of the j th neighbor being $S_{(j),i}^X$.

b) w^X -subproblem: The minimization of (2) with respect to the \mathbf{w}^X can be reformulated as

$$\min_{\mathbf{w}^X} \sum_{m=1}^M w_m \sum_{i,j=1}^N \text{dist}_{m,i,j}^X S_{i,j}^X, \quad \text{s.t. } \sum_{m=1}^M (w_m^X)^\eta = 1, \quad w_m^X \geq 0. \quad (12)$$

By denoting $g_m = \sum_{i,j=1}^N \text{dist}_{m,i,j}^X S_{i,j}^X$, the Lagrangian function of (12) is

$$\mathcal{L}(\mathbf{w}^X, \xi, \boldsymbol{\varsigma}) = \sum_{m=1}^M w_m^X g_m - \xi \left(\sum_{m=1}^M (w_m^X)^\eta - 1 \right) - \boldsymbol{\varsigma}^T w_m \quad (13)$$

where $\xi \in \mathbb{R}$ and $\boldsymbol{\varsigma} \in \mathbb{R}^M$ are two Lagrangian multipliers. With the KKT condition, the closed-form solution of \mathbf{w}^X is given by

$$w_m^X = g_m^{\frac{1}{\eta-1}} \left(\sum_{m=1}^M g_m^{\frac{\eta}{\eta-1}} \right)^{-\frac{1}{\eta}}. \quad (14)$$

From (14), we can find that the feature with smaller distance contribution (g_m) will get larger weight coefficient. We can also observe that when $\eta \rightarrow 1$, we will assign 1 to the weight factor of the feature whose g_m value is the smallest and assign 0 to the weights of other features. By tuning η between (0, 1) we can reach a balance between all the features.

3) k -Selection: It can be found from (6) and (11) that the learned probabilistic graph \mathbf{S}^X is a KNN-type graph, whose number k_i of NNs for each vertex is controlled by the parameter α_i . Obviously, a very small k_i will underconnect the graph and make the graph not robust enough, whereas a very large k_i tends to overconnect the graph and leads to confusion. Here, we use a k -selection strategy proposed in [44] with the goal that “each superpixel is connected to as many truly similar superpixels as possible,” which is an in-degree-based strategy with the following steps.

Step 1: Set $k_{\max} = \lceil \sqrt{N} \rceil$ and $k_{\min} = \lceil \sqrt{N}/10 \rceil$ with $\lceil \cdot \rceil$ denoting the rounding up operation, and find the k_{\max} NNs of each vertex. *Step 2:* Calculate the in-degree $di(\tilde{\mathbf{X}}_i)$ for the i th vertex, that is, the number of times that the i th vertex occurs among the k_{\max} NNs of all vertexes. *Step 3:* Set $k_i = \min\{k_{\max}, \max\{di(\tilde{\mathbf{X}}_i), k_{\min}\}\}$ for the i th vertex.

With this k -selection strategy, we can select a larger k_i for the superpixel that represents more-dominant occupied object (has more truly similar superpixels) and select a smaller k_i for the superpixel that represents less-dominant occupied object (has fewer truly similar superpixels).

The construction process of AG \mathbf{S}^X is reported as Algorithm 1, where the algorithm terminates when the maximal number of iterations is reached or the relative difference between two iteration results $\epsilon^{[t+1]} < \epsilon^{[0]}$ with $\epsilon^{[t+1]} = (\|\mathbf{S}^{X[t+1]} - \mathbf{S}^{X[t]}\|_F)/(\|\mathbf{S}^{X[t+1]}\|_F)$. We can find that it finds the truly similar neighbors for each vertex with two methods: adaptive k -selection strategy and adaptive-weighted distance metric.

Algorithm 1 AG Learning

Input: The feature matrix \mathbf{X} , parameter $\eta \in (0, 1)$.
Initialize: Calculate the distance matrix $dist^X$, adaptively select k_i , and set $\mathbf{w}^X = \mathbf{1}_M$.
Repeat:

- 1: Calculate the weighted distance vector \mathbf{d}_i^X and sort \mathbf{d}_i^X .
- 2: Calculate the \mathbf{S}_i^X through (11).
- 3: Calculate the g_m and update the \mathbf{w}^X through (14).

Output: The learned probability matrix \mathbf{S}^X and weighting vector \mathbf{w}^X .

IV. STRUCTURE CYCLE CONSISTENCY-BASED IMAGE REGRESSION

The main goal of the image regression-based heterogeneous CD is to map the image from one domain to the domain of the other image. Because the superpixel is used as the basic unit and the features are extracted to represent the superpixel, then we need to find the regression function between the feature matrices. Defining the mapping function from the domain of \mathcal{X} to the domain of \mathcal{Y} as $\mathcal{M} : \mathcal{X} \rightarrow \mathcal{Y}$ and the regression feature matrix of \mathbf{X} in the domain of \mathcal{Y} as \mathbf{Y}' , then we have $\mathbf{Y}' = \mathcal{M}(\mathbf{X})$. Similarly, we define the opposite mapping function from the domain of \mathcal{Y} to the domain of \mathcal{X} as: $\mathcal{G} : \mathcal{Y} \rightarrow \mathcal{X}$, and the regression feature matrix of \mathbf{Y} in the domain of \mathcal{X} as \mathbf{X}' ; then, we have $\mathbf{X}' = \mathcal{G}(\mathbf{Y})$. If we define \mathcal{F}^{-1} as the operators for extracting pixel values from features, e.g., directly treating the mean feature as the value of each pixel within the superpixel, then we have the regression function between heterogeneous images as $\mathcal{F}^{-1}\mathcal{M}\mathcal{F}$ and $\mathcal{F}^{-1}\mathcal{G}\mathcal{F}$ with

$$\begin{aligned} \tilde{\mathbf{Y}}' &= \mathcal{F}^{-1}(\mathbf{Y}') = \mathcal{F}^{-1}(\mathcal{M}(\mathbf{X})) = \mathcal{F}^{-1}(\mathcal{M}(\mathcal{F}(\tilde{\mathbf{X}}))) \\ \tilde{\mathbf{X}}' &= \mathcal{F}^{-1}(\mathbf{X}') = \mathcal{F}^{-1}(\mathcal{G}(\mathbf{Y})) = \mathcal{F}^{-1}(\mathcal{G}(\mathcal{F}(\tilde{\mathbf{Y}}))) \end{aligned} \quad (15)$$

where $\tilde{\mathbf{Y}}'$ and $\tilde{\mathbf{X}}'$ are the translated images. Next, we need to use the structure consistency property between heterogeneous images to complete the regression functions of \mathcal{M} and \mathcal{G} by using three types of regularization terms. Fig. 2 shows the structure cycle consistency-based image regression.

A. Forward Transformation

The forward transformation-based regularization (FTR) is used to constrain the regression image and the original image to have the same structure, i.e., the same similarity relationships between superpixels. Specifically, if superpixels in the preevent image ($\tilde{\mathbf{X}}_i$ and $\tilde{\mathbf{X}}_j$) are very similar (representing the same kind of objects), then the superpixels in the transformed image ($\tilde{\mathbf{Y}}'_i$ and $\tilde{\mathbf{Y}}'_j$) corresponding to this superpixel pair should

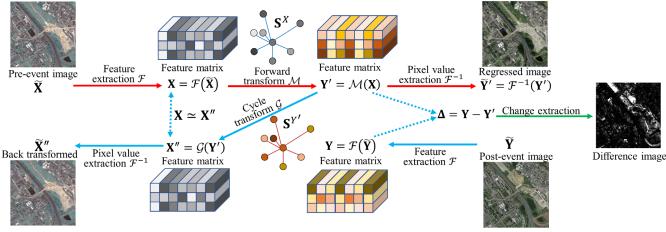


Fig. 2. Illustration of structure cycle consistency-based image regression.

also be very similar (with small distance). Since the structure information (i.e., similarity relationships) of the image $\tilde{\mathbf{X}}$ is characterized by the AG \mathbf{S}^X , then we have the following FTR:

$$\min \sum_{m=1}^M w_m^{Y'} \sum_{i,j=1}^N \text{dist}_{m,i,j}^{Y'} S_{i,j}^X + \beta \|\mathbf{S}^X - \mathbf{S}^{Y'}\|_F^2 \quad (16)$$

where $\text{dist}_{m,i,j}^{Y'} = \|\mathbf{Y}_i^{(m)} - \mathbf{Y}_j^{(m)}\|_2^2$ is the m th feature distance between $\tilde{\mathbf{Y}}_i$ and $\tilde{\mathbf{Y}}_j$, $\beta > 0$ is a tuning parameter, and $\mathbf{S}^{Y'}$ is the structure graph of translated image $\tilde{\mathbf{Y}}'$ that can be learned by using a similar model of (2) as

$$\begin{aligned} & \min_{\mathbf{S}^{Y'}, \mathbf{w}^{Y'}} \sum_{i,j=1}^N \left(\sum_{m=1}^M w_m^{Y'} \text{dist}_{m,i,j}^{Y'} S_{i,j}^{Y'} \right) + \alpha_i \left(S_{i,j}^{Y'} \right)^2 \\ & \text{s.t. } \sum_{i=1}^N S_{i,j}^{Y'} = 1, \quad S_{i,j}^{Y'} \geq 0, \quad \sum_{m=1}^M (w_m^{Y'})^\eta = 1, \quad w_m^{Y'} \geq 0 \end{aligned} \quad (17)$$

where $w_m^{Y'}$ is the weight for the m th feature that is similar as w_m^X in the AG learning model of (2). FTR makes the regressed superpixels corresponding to the instances connected in the edge of \mathbf{S}^X be similar to each other.

Denote the degree matrix $\mathbf{D}_X \in \mathbb{R}^{N \times N}$ of graph \mathbf{S}^X as diagonal matrix with the i th entry being the summation of all the probabilities related to $\tilde{\mathbf{X}}_i$, i.e., $\sum_{j=1}^N (S_{i,j}^X + S_{j,i}^X)/2$, and denote the graph Laplacian matrix as $\mathbf{L}_X = \mathbf{D}_X - (\mathbf{S}^X + (\mathbf{S}^X)^T)/2$; then, we have

$$\sum_{i,j=1}^N \text{dist}_{m,i,j}^{Y'} S_{i,j}^X = 2 \text{Tr} \left(\mathbf{Y}'^{(m)} \mathbf{L}_X \left(\mathbf{Y}'^{(m)} \right)^T \right). \quad (18)$$

B. Cycle Transformation

The cycle transformation-based regularization (CTR) is inspired by the cycle GAN [30], where it pointed out that domain translations should comply with the principle of cycle consistency, that is, the regressed data can be transformed back to the original domain as

$$\mathbf{X}'' = \mathcal{G}(\mathbf{Y}') = \mathcal{G}(\mathcal{M}(\mathbf{X})) \simeq \mathbf{X}. \quad (19)$$

Therefore, we propose the following cycle consistency-based CTR as:

$$\min \sum_{m=1}^M w_m^X \sum_{i,j=1}^N \text{dist}_{m,i,j}^{X''} S_{i,j}^{Y'} + \gamma \sum_{m=1}^M \|\mathbf{X}^{(m)} - \mathbf{X}''^{(m)}\|_F^2 \quad (20)$$

where $\text{dist}_{m,i,j}^{X''} = \|\mathbf{X}_i^{(m)} - \mathbf{X}_j^{(m)}\|_2^2$ and $\gamma > 0$ is a tuning parameter. In the CTR, the first term is based on the structure consistency between the transformed back image and the

transformed image (i.e., the constraint of $\mathbf{X}'' = \mathcal{G}(\mathbf{Y}')$), and the second term requires that the transformed back data are very similar to the original data (i.e., the constraint of $\mathbf{X}'' \simeq \mathbf{X}$).

By defining the graph Laplacian matrix as $\mathbf{L}_{Y'} = \mathbf{D}_{Y'} - (\mathbf{S}^{Y'} + ((\mathbf{S}^{Y'})^T)/2)$ with $\mathbf{D}_{Y'}$ being the degree matrix of graph $\mathbf{S}^{Y'}$, CTR can be rewritten as

$$\min \sum_{m=1}^M 2w_m^X \text{Tr} \left(\mathbf{X}''^{(m)} \mathbf{L}_{Y'} \left(\mathbf{X}''^{(m)} \right)^T \right) + \gamma \left\| \mathbf{X}^{(m)} - \mathbf{X}''^{(m)} \right\|_F^2. \quad (21)$$

C. Sparse Regularization

One of the challenges in unsupervised image regression or heterogeneous CD problem is to avoid that changed pixels affect the transforming functions. To reduce the influence of changed pixels on the regression process, we decompose postevent image into a regressed image and a changed image, and then, we have $\mathbf{Y} = \mathbf{Y}' - \Delta$, where $\Delta \in \mathbb{R}^{M \times C_Y \times N}$ is the changed feature matrix. Based on the fact that only a small part of the area changed and most areas remain unchanged during the event in the practical CD problem, we have the following prior sparsity-based regularization (PSR) as:

$$\min \lambda \sum_{m=1}^M \|\Delta^{(m)}\|_{2,1}, \quad \text{s.t. } \mathbf{Y}^{(m)} = \mathbf{Y}'^{(m)} + \Delta^{(m)} \quad (22)$$

where $\lambda > 0$ is a penalty parameter. The $\ell_{2,1}$ -norm $\|\Delta^{(m)}\|_{2,1}$ is defined as $\|\Delta^{(m)}\|_{2,1} = \sum_{i=1}^N \|\Delta_i^{(m)}\|_2$, which is a convex relaxation of the original $\ell_{2,0}$ -norm of $\|\Delta^{(m)}\|_{2,0}$ that equals the number of nonzero columns in $\Delta^{(m)}$.

D. Image Regression Model

By combining the above regularization terms of FTR (16), CTR (20), PSR (22), and the graph learning model (17), we can obtain the structure cycle consistency-based image regression model as

$$\begin{aligned} & \min_{\mathbf{X}'', \mathbf{Y}', \Delta, \mathbf{S}^{Y'}, \mathbf{w}^{Y'}} \sum_{m=1}^M 2w_m^{Y'} \text{Tr} \left(\mathbf{Y}'^{(m)} (\mathbf{L}_X + \mathbf{L}_{Y'}) \left(\mathbf{Y}'^{(m)} \right)^T \right) \\ & + 2w_m^X \text{Tr} \left(\mathbf{X}''^{(m)} \mathbf{L}_{Y'} \left(\mathbf{X}''^{(m)} \right)^T \right) \\ & + \gamma \left\| \mathbf{X}^{(m)} - \mathbf{X}''^{(m)} \right\|_F^2 + \lambda \|\Delta^{(m)}\|_{2,1} \\ & + \sum_{i=1}^N \alpha_i \left\| \mathbf{S}_i^{Y'} \right\|_2^2 + \beta \left\| \mathbf{S}^X - \mathbf{S}^{Y'} \right\|_F^2 \\ & \text{s.t. } \sum_{i=1}^N S_{i,j}^{Y'} = 1, \quad S_{i,j}^{Y'} \geq 0, \quad \sum_{m=1}^M (w_m^{Y'})^\eta = 1, \quad w_m^{Y'} \geq 0 \\ & \mathbf{Y}'^{(m)} = \mathbf{Y}^{(m)} + \Delta^{(m)}. \end{aligned} \quad (23)$$

With the regression model (23), it can be found that the regressed \mathbf{Y} and \mathbf{Y}' are in the same domain, for two reasons. First, \mathbf{Y}' is separated from \mathbf{Y} and only a few of them are different (Δ is column sparse), i.e., $\mathbf{Y}'_i = \mathbf{Y}_i$ ideally holds for most $i = 1, \dots, N$. Second, for the changed \mathbf{Y}'_i , the model

constrains it to be similar to its neighborhood of $\{\mathbf{Y}'_i | S_{i,j}^X \neq 0, i = 1, \dots, N\}$, which prevents anomalous \mathbf{Y}'_j .

By using the ADMM and defining the objective function of (23) as $\Theta(\mathbf{X}'', \mathbf{Y}', \Delta, \mathbf{S}^{Y'}, \mathbf{w}^{Y'})$, we rewrite the model (23) as the minimization of following augmented Lagrangian function:

$$\begin{aligned} & \mathcal{L}(\mathbf{S}^{Y'}, \mathbf{X}'', \mathbf{Y}', \Delta, \mathbf{w}^{Y'}, \mathbf{P}) \\ &= \Theta + \sum_{m=1}^M \text{Tr} \left((\mathbf{Y}'^{(m)} - \mathbf{Y}^{(m)} - \Delta^{(m)})^T \mathbf{P}^{(m)} \right) \\ & \quad + \sum_{m=1}^M \frac{\mu}{2} \left\| \mathbf{Y}'^{(m)} - \mathbf{Y}^{(m)} - \Delta^{(m)} \right\|_F^2 \\ \text{s.t. } & \sum_{i=1}^N S_{i,j}^{Y'} = 1, \quad S_{i,j}^{Y'} \geq 0, \quad \sum_{m=1}^M (w_m^{Y'})^\eta = 1, \quad w_m^{Y'} \geq 0 \end{aligned} \quad (24)$$

where $\mathbf{P}^{(m)} \in \mathbb{R}^{C_Y \times N}$ is a Lagrangian multiplier and $\mu > 0$ is a penalty parameter. The minimization problem of (24) can be divided into the following subproblems.

1) $S^{Y'}$ -Subproblem: By defining $\mathbf{d}_i^{Y'} = \sum_{m=1}^M (w_m^{Y'} \text{dist}_{m,i}^{Y'} + w_m^X \text{dist}_{m,i}^{X''}) - 2\beta S_i^X$, the minimization of (24) with respect to $\mathbf{S}^{Y'}$ can be reformulated as

$$\min_{\mathbf{1}_N^T \mathbf{S}_i^{Y'} = 1, S_{i,j}^{Y'} \geq 0} \left\| \mathbf{S}_i^{Y'} + \frac{1}{2(\alpha_i + \beta)} \mathbf{d}_i^{Y'} \right\|_2^2. \quad (25)$$

Similar to the process of solving \mathbf{S}^X in (11), the solution for $\mathbf{S}^{Y'}$ is

$$S_{(j),i}^{Y'} = \begin{cases} \frac{d_{(k_i+1),i}^{Y'} - d_{(j),i}^{Y'}}{k_i d_{(k_i+1),i}^{Y'} - \sum_{h=1}^{k_i} d_{(h),i}^{Y'}}, & j \leq k_i \\ 0, & j > k_i \end{cases} \quad (26)$$

where k_i of $\mathbf{S}^{Y'}$ for the i th vertex is same as k_i of \mathbf{S}^X with the k -selection strategy in Section III-B3.

2) X'' -Subproblem: The minimization of (24) with respect to \mathbf{X}'' can be reformulated as

$$\min_{\mathbf{X}''} \sum_{m=1}^M 2w_m^X \text{Tr} \left(\mathbf{X}''^{(m)} \mathbf{L}_{Y'} (\mathbf{X}''^{(m)})^T \right) + \gamma \left\| \mathbf{X}^{(m)} - \mathbf{X}''^{(m)} \right\|_F^2. \quad (27)$$

Then, we have the optimal \mathbf{X}'' as

$$\mathbf{X}''^{(m)} = 2\gamma \mathbf{X}^{(m)} (4w_m^X \mathbf{L}_{Y'} + 2\gamma \mathbf{I}_N)^{-1} \quad (28)$$

where $\mathbf{I}_N \in \mathbb{R}^{N \times N}$ represents an identity matrix.

3) Y' -Subproblem: The minimization of (24) with respect to \mathbf{Y}' can be reformulated as

$$\begin{aligned} & \min_{\mathbf{Y}'} \sum_{m=1}^M 2w_m^{Y'} \text{Tr} \left(\mathbf{Y}'^{(m)} (\mathbf{L}_X + \mathbf{L}_{Y'}) (\mathbf{Y}'^{(m)})^T \right) \\ & \quad + \text{Tr} \left((\mathbf{Y}'^{(m)})^T \mathbf{P}^{(m)} \right) + \frac{\mu}{2} \left\| \mathbf{Y}'^{(m)} - \mathbf{Y}^{(m)} - \Delta^{(m)} \right\|_F^2. \end{aligned} \quad (29)$$

The \mathbf{Y}' -subproblem can be solved by taking the first-order derivative of the objective function to zero, and then, we have

$$\mathbf{Y}'^{(m)} = (\mu \mathbf{Y}^{(m)} + \mu \Delta^{(m)} - \mathbf{P}^{(m)}) \left(4w_m^{Y'} (\mathbf{L}_X + \mathbf{L}_{Y'}) + \mu \mathbf{I}_N \right)^{-1}. \quad (30)$$

4) Δ -Subproblem: The minimization of (24) with respect to Δ can be rewritten as

$$\min_{\Delta} \sum_{m=1}^M \left\| \Delta^{(m)} \right\|_{2,1} + \frac{\mu}{2\lambda} \left\| \Delta^{(m)} + \mathbf{Y}^{(m)} - \mathbf{Y}'^{(m)} - \frac{\mathbf{P}^{(m)}}{\mu} \right\|_F^2. \quad (31)$$

By defining $\mathbf{Q}^{(m)} = -\mathbf{Y}^{(m)} + \mathbf{Y}'^{(m)} + (\mathbf{P}^{(m)}/\mu)$ and the proximal operator $\text{prox}_{af}(\mathbf{b}) = \arg \min_{\mathbf{x}} f(\mathbf{x}) + (1/2a) \|\mathbf{x} - \mathbf{b}\|_F^2$, the closed-form solution of (31) can be obtained by $\Delta^{(m)} = \text{prox}_{(\lambda/\mu)\|\cdot\|_{2,1}}(\mathbf{Q}^{(m)})$ as in [45] with

$$\Delta_i^{(m)} = \max \left\{ \left\| \mathbf{Q}_i^{(m)} \right\|_2 - \frac{\lambda}{\mu} \right\} \frac{\mathbf{Q}_i^{(m)}}{\left\| \mathbf{Q}_i^{(m)} \right\|_2} \quad (32)$$

where we follow the convention $0 \cdot (0/0) = 0$.

5) $w^{Y'}$ -Subproblem: The minimization of (24) with respect to $\mathbf{w}^{Y'}$ can be reformulated as

$$\begin{aligned} & \min_{\mathbf{w}^{Y'}} \sum_{m=1}^M w_m^{Y'} \sum_{i,j}^N (S_{i,j}^X + S_{i,j}^{Y'}) \text{dist}_{m,i,j}^{Y'} \\ \text{s.t. } & \sum_{m=1}^M (w_m^{Y'})^\eta = 1, \quad w_m^{Y'} \geq 0. \end{aligned} \quad (33)$$

By denoting $\rho_m = \sum_{i,j}^N (S_{i,j}^X + S_{i,j}^{Y'}) \text{dist}_{m,i,j}^{Y'}$, the $\mathbf{w}^{Y'}$ -subproblem can be solved similarly as \mathbf{w}^X -subproblem of (12), and then, we have

$$w_m^{Y'} = \rho_m^{\frac{1}{\eta-1}} \left(\sum_{m=1}^M \rho_m^{\frac{\eta}{\eta-1}} \right)^{-\frac{1}{\eta}}. \quad (34)$$

Finally, the Lagrangian multiplier \mathbf{P} can be updated by

$$\mathbf{P}^{(m)} \leftarrow \mathbf{P}^{(m)} + \mu \left(\mathbf{Y}'^{(m)} - \mathbf{Y}^{(m)} - \Delta^{(m)} \right). \quad (35)$$

The procedure of solving the problem (23) is summarized as Algorithm 2. The stopping criterion of Algorithm 2 is that the maximum number of iterations is reached or the relative difference between two iteration results $\varepsilon^{[t+1]} = \sum_{m=1}^M \|\Delta^{(m)[t+1]} - \Delta^{(m)[t]}\|_F / \sum_{m=1}^M \|\Delta^{(m)[t+1]}\|_F$ is less than the tolerance threshold $\varepsilon^{[0]}$, which means that there is no longer any appreciate updates in the iteration and the algorithm runs into convergence.

E. Change Extraction

Once the regressed feature matrix \mathbf{Y}' and changed feature matrix Δ are output from Algorithm 2, the regression image can be obtained by using $\tilde{\mathbf{Y}'} = \mathcal{F}^{-1}(\mathbf{Y}')$, and the DI can be calculated by

$$\text{DI}(i, j) = \sum_{m=1}^M \left\| \Delta_m^{(m)} \right\|_2^2, \quad (i, j) \in \Lambda_n, \quad n = 1, \dots, N. \quad (36)$$

After the DI is obtained, the binary CM solution can be regarded as an image segmentation problem, which can be solved by thresholding methods or clustering methods as used in homogeneous CD. Since the main purpose of this work is to

Algorithm 2 Structure Cycle Consistency-Based Image Regression

Input: The matrices of \mathbf{X} , \mathbf{Y} , \mathbf{S}^X and \mathbf{w}^X , parameters $\beta, \gamma, \lambda > 0$.

Initialize: Set $\mathbf{X}'' = \mathbf{X}$, $\mathbf{Y}' = \mathbf{Y}$, $\mathbf{w}^{Y'} = \mathbf{w}^X$, and $\mathbf{P} = \Delta = \mathbf{0}$.

Repeat:

- 1: Update $\mathbf{S}^{Y'}$ through (26) by computing and sorting $\mathbf{d}_i^{Y'}$.
- 2: Update \mathbf{X}'' through (28) by computing $\mathbf{L}_{Y'}$.
- 3: Update \mathbf{Y}' through (30).
- 4: Update Δ through (32) by computing \mathbf{Q} .
- 5: Update $\mathbf{w}^{Y'}$ through (34) by computing ρ_m .
- 6: Update the Lagrangian multiplier through (35).

Until stopping criterion is met.

Output: The regressed feature matrix \mathbf{Y}' and changed feature matrix Δ .

propose the AGSCC-based image regression method for heterogeneous CD, a complex segmentation method for computing final CM is outside the focus of this article, which may also conceal the contribution of this article. Therefore, we directly use the superpixel-based segmentation method proposed in SCASC [39] to divide DI into changed and unchanged classes, which exploits spatial context information and change information of DI with a Markov random field (MRF) model that can be solved by using the min-cut/max-flow algorithm [46]. The only hyperparameter that needs to be manually set in this MRF-based segmentation method is the balanced parameter, which is fixed to 0.05 according to SCASC [39].

V. EXPERIMENTAL RESULTS AND DISCUSSION

A. Heterogeneous Datasets and Evaluation Metrics

Six heterogeneous datasets are presented to evaluate the proposed AGSCC as listed in Table I, which contains different types of heterogeneity: multisensor image pairs (e.g., #1, #2, and #3) and multsource image pairs (e.g., #4, #5, and #6). These datasets reflect quite different CD conditions: different resolution levels (varying from 0.52 to 30 m), different image sizes (varying from 300 to 4135 pixels in length or width), and different change events (lake overflow, flooding, and construction), which can evaluate the generalizability and robustness of the algorithm.

Two types of evaluation metrics are employed. First, the DI can be evaluated by the empirical receiver operating characteristics (ROC) curve and the precision-recall (PR) curve, and the corresponding areas under ROC curve (AUR) and PR curve (AUP) are used as the quantitative criteria, respectively. Second, the final CM can be evaluated by the overall accuracy (OA), F1-measure (Fm), and Kappa coefficient (Kc), which are computed as: $OA = (TP + TN)/(TP + TN + FP + FN)$, $Fm = (2TP)/(2TP + FP + FN)$, and $Kc = (OA - PRE)/(1 - PRE)$ with

$$PRE = \frac{(TP + FN)(TP + FP) + (TN + FP)(TN + FN)}{(TP + TN + FP + FN)^2} \quad (37)$$

where TP, FP, TN, and FN represent the true positives, false positives, true negatives, and false negatives, respectively.

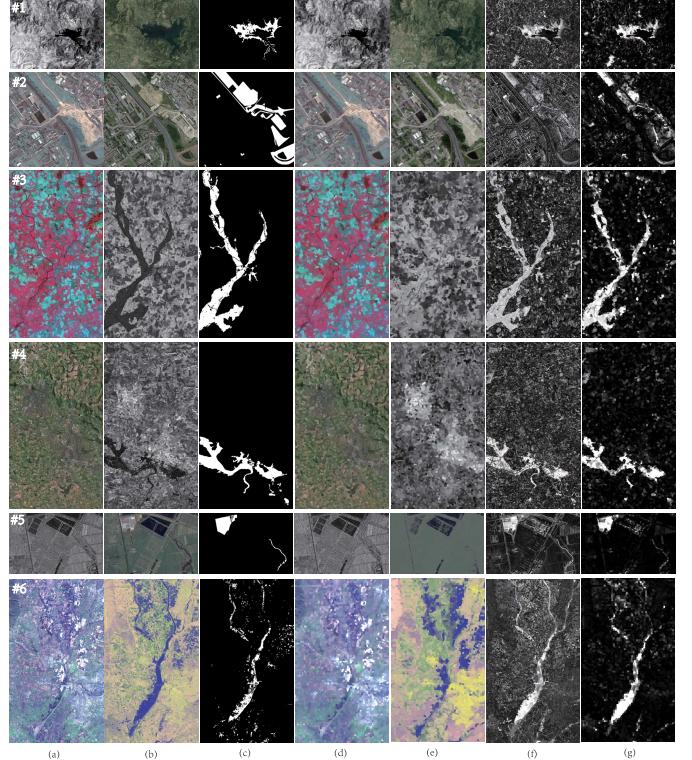


Fig. 3. Heterogeneous datasets, regression images, and DIs. From top to bottom, they correspond to Datasets #1–#6. (a) Preevent image $\tilde{\mathbf{X}}$, (b) postevent image $\tilde{\mathbf{Y}}$, (c) ground truth, (d) back transformed image $\tilde{\mathbf{X}}''$, (e) regression images $\tilde{\mathbf{Y}}'$, (f) DI between $\tilde{\mathbf{Y}}$ and $\tilde{\mathbf{Y}}'$, and (g) DI calculated by using changed feature matrix Δ with (36).

Combining the comprehensive metrics (OA, Kc and Fm) with the individual metrics (TP, FP, TN, and FN marked in different colors in the qualitative results) allows us to better assess the CM.

For all the experiments of AGSCC, we set $N = 5000$ for the superpixel segmentation and extract the mean, variance, and median values of each band as the features of each superpixel; set $\eta = 0.5$ for Algorithm 1; and set $\beta = \gamma = 5 \times \|\mathbf{w}^X\|_1$ and $\lambda = 0.1 \times \|\mathbf{w}^X\|_1$ for Algorithm 2. The impact of these parameters will be analyzed in detail in Section V-D in the Supplementary Material.

B. Regression Image and DI Performance

In Fig. 3(d) and (e), the back transformed images $\tilde{\mathbf{X}}''$ and the regression images $\tilde{\mathbf{Y}}'$ of the preevent image are shown, which are computed by $\tilde{\mathbf{X}}'' = \mathcal{F}^{-1}(\mathbf{X}'')$ and $\tilde{\mathbf{Y}}' = \mathcal{F}^{-1}(\mathbf{Y}')$ with \mathcal{F}^{-1} being the operator of extracting the mean features from the feature matrix, that is, we set the pixel value within each superpixel of $\tilde{\mathbf{X}}''$ (or $\tilde{\mathbf{Y}}'$) to the mean feature of each band of this superpixel in \mathbf{X}'' (or \mathbf{Y}'). From Fig. 3(a) and (d), we can find that the back transformed image $\tilde{\mathbf{X}}''$ and the original preevent image $\tilde{\mathbf{X}}$ are very similar, except that $\tilde{\mathbf{X}}''$ is blurred due to the use of mean features. By comparing the translated image $\tilde{\mathbf{Y}}'$ in Fig. 3(e), the original preevent image $\tilde{\mathbf{X}}$ in Fig. 3(a), and the postevent image $\tilde{\mathbf{Y}}$ in Fig. 3(b), we can find that the structures of $\tilde{\mathbf{Y}}'$ and $\tilde{\mathbf{X}}$ are consistent in both the changed and unchanged regions, and $\tilde{\mathbf{Y}}'$ and $\tilde{\mathbf{Y}}$ have the same

TABLE I
DESCRIPTION OF THE SIX HETEROGENEOUS DATASETS

Dataset	Sensor	Size (pixels)	Date	Location	Event (& Spatial resolution)
#1	Landsat-5/Google Earth	300 × 412 × 1(3)	Sept. 1995 - July 1996	Sardinia, Italy	Lake expansion (30m.)
#2	Pleiades/WorldView2	2000 × 2000 × 3(3)	May 2012 - July 2013	Toulouse, France	Construction (0.52m.)
#3	Spot/NDVI	990 × 554 × 3(1)	1999 - 2000	Gloucester, England	Flooding ($\approx 25m.$)
#4	Quick Bird 2/TerraSAR-X	4135 × 2325 × 3(1)	July 2006 - July 2007	Gloucester, England	Flooding (0.65m)
#5	Radarsat-2/Google Earth	593 × 921 × 1(3)	June 2008 - Sept. 2012	Shuguang Village, China	Building construction (8m.)
#6	Landsat-8/Sentinel-1A	875 × 500 × 11(3)	Jan. 2017 - Feb. 2017	Sutter County, USA	Flooding ($\approx 15m.$)

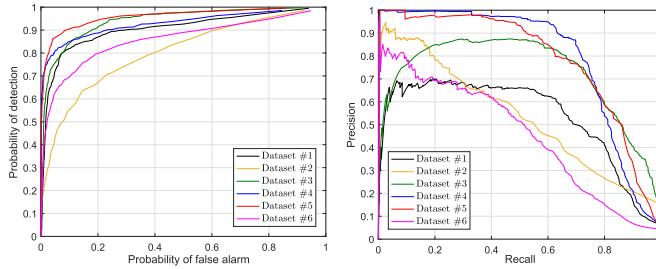


Fig. 4. ROC (left) and PR (right) curves of AGSCC generated DIs.

image style (i.e., their statistical properties are similar). This means that the proposed structure consistency-based AGSCC can well translate the heterogeneous images from different domains.

Fig. 3(f) shows the DI between \tilde{Y} and \tilde{Y}' , which is computed by directly using the differencing operator as $DI(i, j) = (\sum_{c=1}^{C_Y} (\tilde{y}(i, j, c) - \tilde{y}'(i, j, c))^2)^{1/2}$. Fig. 3(g) shows the DI calculated by using changed feature matrix Δ with (36). From Fig. 3(f) and (g), we can see that both DIs are able to well distinguish the changed and unchanged parts, which demonstrates the effectiveness of Algorithms 1 and 2 in the structure representation and image transformation with the structure cycle consistency, respectively. We can also find that the DIs with Δ in Fig. 3(g) are sparse by using the PSR. Therefore, it is able to obtain a satisfactory CD performance by directly segmenting the DI with a simple thresholding method (such as the Otsu [47]) or clustering method (such as K-means clustering [48] and fuzzy c-means clustering [49]), which is also confirmed by the ROC and PR curves in Fig. 4. The AUR of the ROC curves and the AUP of PR curves are reported in Table II. Focusing on Fig. 4 and Table II, the DIs generated by AGSCC on Datasets #2 and #6 are not as good as others. For Datasets #2 and #6, they contain more types of ground objects than other datasets, such as buildings, grass, trees, roads, and pitches of Dataset #2, and rivers, farmland, roads, buildings, forests, and mountains of Dataset #6. At the same time, the proportion of these ground objects is quite uneven and the resolution of Dataset #2 is very high (0.52 m), as shown in Fig. 3. Therefore, due to the complex image structure of Datasets #2 and #6, it is difficult to accurately capture the structure information and thus accomplish a good structure-based image regression. In contrast, the DIs of Datasets #3, #4, and #5 can obtain very high metrics, for example, the AUPs are 0.719, 0.797, and 0.787, respectively. In addition, we believe that the proposed method can be further expanded by incorporating

TABLE II
AUR AND AUP OF DIS GENERATED BY AGSCC ON THE HETEROGENEOUS DATASETS

Measures	Datasets					
	#1	#2	#3	#4	#5	#6
AUR	0.905	0.802	0.940	0.928	0.959	0.858
AUP	0.532	0.534	0.719	0.797	0.787	0.452

with deep learning-based methods. For example, AGSCC can be associated with some homogeneous CD methods after acquiring the regression image [7], [8]; and AGSCC can also provide assistance to some deep learning-based CD methods, such as supporting the training process [29] or constructing high confidence pseudo-training sets [22], [34].

C. CM Performance

To verify the effectiveness of the proposed method, we select the recently proposed M3CD¹ [50], NPSG² [36], ALSC [43], IRG-McS (with the similarity criterion)³ [44], FPMS⁴ [37], SCASC⁵ [39], PSGM [38], and GIR-MRF⁶ [40] for comparison. The default parameters are used in their codes, which are also consistent with their original papers. Fig. 5 shows the binary CMs of different methods on all the evaluated datasets, and Table III lists the corresponding criteria of OA, Kc, and Fm. From Fig. 5, we can find that the AGSCC can well detect the changes between most heterogeneous images with relatively small FP and FN. At the same time, by comparing the quantitative measures of different methods in Table III, it can be found that some methods do not perform robustly enough and their performance degrades considerably on some complex scenarios, such as M3CD on Dataset #6 (Kc is 0.021), NPSG on Dataset #4 (Kc is 0.413), and FPMS and ALSC on Dataset #2 (Kc are 0.215 and 0.312, respectively). In contrast, the proposed AGSCC can obtain robust detection results on different datasets under different types of CD conditions. The average OA, Kc, and Fm of AGSCC on all the evaluated datasets are about 0.955, 0.661, and 0.684, respectively, which are higher than other comparison methods. For example, the average Kc of AGSCC is 1.5% higher than

¹M3CD is kindly available at <http://www-labs.iro.umontreal.ca/~mignotte>

²NPSG is available at <https://github.com/yulisun/NPSG>

³IRG-McS is available at <https://github.com/yulisun/IRG-McS>

⁴FPMS is kindly available at <http://www-labs.iro.umontreal.ca/~mignotte>

⁵SCASC is available at <https://github.com/yulisun/SCASC>

⁶GIR-MRF is available at <https://github.com/yulisun/GIR-MRF>

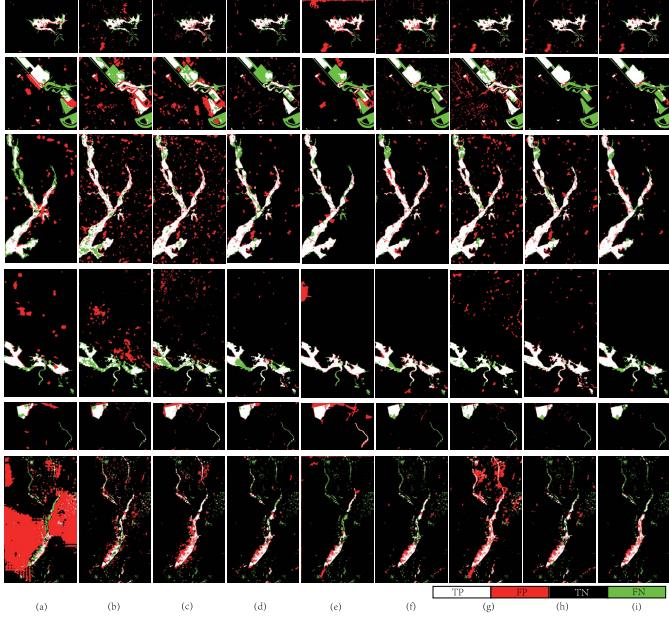


Fig. 5. Binary CMs of different methods on heterogeneous datasets. From top to bottom, they correspond to Datasets #1–#6. Binary CMs generated by (a) M3CD, (b) NPSG, (c) ALSC, (d) IRG-McS (with the similarity criterion), (e) FPMS, (f) SCASC, (g) PSGM, (h) GIR-MRF, and (i) AGSCC. In the binary CM, white indicates TP, red indicates FP, black indicates TN, and green indicates FN.

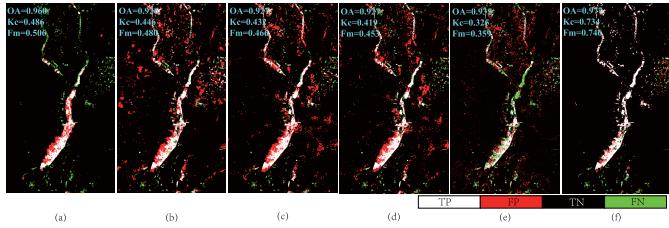


Fig. 6. Binary CMs of different methods on Dataset #6: (a) AGSCC, (b) cGAN, (c) ACE-Net, (d) X-Net, (e) USCDN, and (f) DTCDN. In the binary CM, white indicates TP, red indicates FP, black indicates TN, and green indicates FN.

the second ranked GIR-MRF and 2.4% higher than the image regression-based SCASC. This demonstrates the effectiveness of the AGSCC-based image regression.

We also provide a comparison with some heterogeneous CD methods based on GAN, such as cGAN [28], ACE-Net [29], X-Net [29], USCDN [31], and DTCDN [27]. Fig. 6 shows the CMs generated by these comparison methods on the challenging Dataset #6, where we have marked the evaluation scores of OA, Kc, and Fm. It can be seen from Fig. 6 that DTCDN obtains quite accurate results, much better than other methods. However, DTCDN is a supervised method, which uses 50% of the ground truth as training data, 20% as the validation set, and 30% as the test set in the deep translation network and the CD network. For the rest unsupervised methods, the proposed AGSCC is still very competitive, achieving better detection results with higher OA, Kc, and Fm.

In addition, in order to further evaluate the performance of AGSCC, some other representative and SOTA methods other than M3CD, NPSG, IRG-McS, FPMS, SCASC, and PSGM

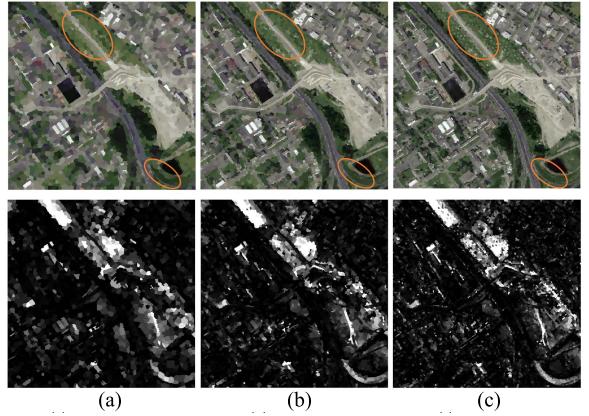


Fig. 7. Regression images and DIs of AGSCC on Dataset #2 with different N 's. (a) $N = 2500$. (b) $N = 5000$. (c) $N = 10000$. From (a) to (c), the AUR of DIs are 0.793, 0.802, and 0.804, respectively, and the AUP of DIs are 0.505, 0.534, and 0.540, respectively.

are selected for comparison as listed in Table IV, including DFR-MT [51], CACFL [23], MDS [52], AFL-DSR [53], RMN [54], PUSM [55], NLPEM [19], DPFL [56], X-Net [29], ACE-Net [29], SCCN [20], LT-FL [57], AMD-IR [26], and SSL [58]. Among these comparison approaches, DFR-MT, CACFL, AFL-DSR, PUSM, DPFL, X-Net, ACE-Net, SCCN, LT-FL, and SSL are deep learning-based methods. For the sake of fairness, we directly quote the results of the corresponding datasets in their original published papers. Because the datasets used in each paper are not identical, Table IV is not aligned. As can be seen in Table IV, the AGSCC can consistently gain better or very competitive accuracy by comparing with these SOTA approaches across different datasets.

D. Discussion

1) Parameter Analysis: The main parameters in AGSCC are: the number of superpixels N ; the feature weight controlling parameter η in the AG learning of Algorithm 1; and the penalty parameters of β , γ , and λ that control the weights of regularization terms in the regression model (23).

Generally, N should be selected according to the image resolution and granularity requirement of CD task. A larger N will make the segmented superpixel smaller, which improves the detection granularity but also increases the computational complexity, as analyzed in the following subsection of complexity analysis. Fig. 7 plots the regressed images and DIs generated by AGSCC on Dataset #2 with $N = 2500$, 5000, and 10000, where we mark some details with the ellipse. We can find that when N is very small (e.g., $N = 2500$), the size of the generated superpixels is very large. In this case, some superpixels may not be internally homogeneous, i.e., they may contain several different kinds of objects. As a result, some details in the regressed image are easy to be ignored, and the block effect of the DI is more obvious, as shown in Fig. 7. On the other hand, a very large N increases the number of nodes in the graph, which in turn increases the complexities of the AG learning (Algorithm 1) and structure cycle consistency-based

TABLE III

QUANTITATIVE MEASURES OF BINARY CMs ON THE HETEROGENEOUS DATASETS. THE HIGHEST AND SECOND-HIGHEST SCORES ARE HIGHLIGHTED IN BOLD AND UNDERLINED, RESPECTIVELY

Methods	Dataset #1			Dataset #2			Dataset #3			Dataset #4			Dataset #5			Dataset #6			Average		
	OA	Kc	Fm																		
M3CD	0.963	0.669	0.689	0.863	0.405	0.481	0.915	0.588	0.636	0.955	0.641	0.665	0.962	0.602	0.622	0.575	0.021	0.077	0.872	0.487	0.528
NPSG	0.947	0.559	0.587	0.830	0.346	0.446	0.902	0.608	0.663	0.929	0.413	0.451	0.975	0.729	0.742	0.941	0.419	0.449	0.921	0.512	0.556
ALSC	0.962	0.698	0.718	0.815	0.312	0.422	0.907	0.641	0.693	0.943	0.515	0.546	0.963	0.669	0.688	0.944	0.470	0.498	0.922	0.551	0.594
IRG-McS	0.971	0.733	0.749	0.881	0.421	0.481	0.942	0.735	0.768	0.972	0.740	0.755	0.978	0.755	0.767	0.958	0.487	0.508	<u>0.950</u>	0.645	0.671
FPMS	0.925	0.552	0.588	0.838	0.215	0.296	0.962	0.816	0.837	0.970	0.771	0.787	0.938	0.569	0.597	0.947	0.329	0.356	0.930	0.542	0.577
SCASC	0.947	0.593	0.621	0.892	0.464	0.516	0.949	0.771	0.800	0.973	<u>0.774</u>	0.789	0.979	0.741	0.751	0.961	0.479	0.500	0.950	0.637	0.663
PSGM	0.961	0.682	0.703	0.857	0.473	0.558	0.922	0.675	0.719	0.959	0.666	0.688	0.977	0.744	0.756	0.908	0.383	0.422	0.931	0.604	0.641
GIR-MRF	0.953	0.620	0.643	<u>0.896</u>	<u>0.484</u>	0.535	0.932	0.719	0.758	<u>0.975</u>	0.801	0.813	0.979	<u>0.772</u>	0.783	0.956	0.482	0.504	0.949	<u>0.646</u>	<u>0.672</u>
AGSCC	0.959	0.658	0.680	0.897	0.490	<u>0.540</u>	0.955	<u>0.791</u>	<u>0.817</u>	0.976	0.766	0.779	0.983	0.773	<u>0.782</u>	<u>0.960</u>	<u>0.486</u>	<u>0.506</u>	0.955	0.661	0.684

TABLE IV

ACCURACY RATE OF CMs GENERATED BY DIFFERENT METHODS ON DIFFERENT DATASETS. THE RESULTS OF THESE COMPARISON METHODS ARE REPORTED BY THEIR ORIGINAL PUBLISHED PAPERS. ITALICIZED AND UNDERLINED MARKS ARE USED FOR DEEP LEARNING-BASED METHODS

Dataset #1	OA
<i>DFR-MT</i> [51]	0.975
<i>CACFL</i> [23]	0.975
AGSCC	0.959
<i>MDS</i> [50]	0.942
<i>AFL-DSR</i> [53]	0.929
<i>RMN</i> [54]	0.847

Dataset #2	OA
AGSCC	0.897
<i>AFL-DSR</i> [53]	0.880
<i>RMN</i> [54]	0.877
<i>PUSM</i> [55]	0.865
<i>NLPEM</i> [19]	0.853

Dataset #3	OA
<i>HPT</i> [25]	0.957-0.964
AGSCC	0.955
<i>AFL-DSR</i> [53]	0.836
<i>MDER</i> [59]	0.818

Dataset #4	OA
AGSCC	0.976
<i>NLPEM</i> [19]	0.949
<i>RMN</i> [54]	0.943
<i>AFL-DSR</i> [53]	0.892

Dataset #5	OA
<i>DPFL</i> [56]	0.987
<i>X-Net</i> [29]	0.984
AGSCC	0.983

Dataset #6	OA
AGSCC	0.960
<i>DPFL</i> [56]	0.945
<i>CACFL</i> [23]	0.979
<i>AMD-IR</i> [26]	0.933
<i>SSL</i> [58]	0.924

image regression (Algorithm 2). In this article, we simply set $N = 5000$ as a compromise choice. In addition, we directly assign the segmentation map Δ of image $\tilde{\mathbf{X}}$ to the image $\tilde{\mathbf{Y}}$ in the preprocessing of AGSCC. Although this operation is very simple, it may introduce inaccurate boundaries of CD. How to better segment the image in the preprocessing of AGSCC, i.e., selecting suitable cosegmentation methods and optimal segmentation scales [60] or fusing multiscale information [7], needs to be further investigated.

To measure the impact of other parameters, the AUR and OA are adopted to evaluate the generated DI and CM, respectively, which can describe the general performance of AGSCC. We take the approach of adjusting the parameter under investigation and fixing the other parameters. In Fig. 8, we vary the parameter η from 0.2 to 0.8 with step 0.05, let $\beta = \gamma$ and set them to $\{1, 2, \dots, 10\} \times \|\mathbf{w}^X\|_1$, and set the parameter λ to $\{0.05, 0.1, \dots, 0.5\} \times \|\mathbf{w}^X\|_1$. From Fig. 8, we can see that AGSCC achieves good results for a fairly large range of the parameters β and γ , which indicates that AGSCC is certain robust to β and γ . In contrast, AGSCC is more sensitive to the parameters η and γ . The larger η is, the more unevenly the weights of \mathbf{w}^X (14) and \mathbf{w}^Y (34) are distributed. In our experiments, we set $\eta = 0.5$ as a compromise choice. For the parameter λ that controls the weights of PSR, its value should be set according to the sparsity level (proportion of the changed regions). Generally, the smaller the changed regions,

the larger λ should be. Based on Fig. 8, we fix $\lambda = 0.1 \times \|\mathbf{w}^X\|_1$ in this article for simplicity.

2) *Complexity Analysis:* The main complexity of AGSCC is concentrated in the AG learning (Algorithm 1) and structure cycle consistency-based image regression (Algorithm 2).

Algorithm 1: in Step 1, calculating the weighted distance matrix \mathbf{d}^X between all the superpixels needs $\mathcal{O}(MC_X N^2/2)$ and sorting all columns of \mathbf{d}^X needs $\mathcal{O}(N^2 \log N)$. In Step 2, calculating the \mathbf{S}^X with the closed form (11) needs $\mathcal{O}(N^2)$. In Step 3, updating \mathbf{w}^X with g_m needs $\mathcal{O}(MN^2)$.

Algorithm 2: In Step 1, updating \mathbf{S}^Y through (26) by computing and sorting \mathbf{d}_i^Y needs $\mathcal{O}(N^2(M(C_X + C_Y)/2 + \log N + 1))$. In Step 2, updating \mathbf{X}'' through (28) requires $\mathcal{O}(MN^3)$ for matrix inversion of $(4w_m^X \mathbf{L}_Y + 2\gamma \mathbf{I}_N)^{-1}$ and $\mathcal{O}(MC_X N^2)$ for matrix multiplication. In Step 3, updating \mathbf{Y}' through (30) also requires $\mathcal{O}(MN^3)$ for matrix inversion of $(4w_m^Y (\mathbf{L}_X + \mathbf{L}_Y) + \mu \mathbf{I}_N)^{-1}$ and $\mathcal{O}(MC_Y N^2)$ for matrix multiplication. In Step 4, updating Δ through (32) requires $\mathcal{O}(MC_Y N)$ with the closed-form proximal operator. In Step 5, updating \mathbf{w}^Y with ρ_m requires $\mathcal{O}(MN^2)$. In Step 6, updating the Lagrangian multiplier through (35) requires $\mathcal{O}(MC_Y N)$ for matrix multiplication.

Although the complexity of AGSCC is very high in the above theoretical analysis, which requires $\mathcal{O}(MN^3)$ for matrix inversion of each iteration in Algorithm 2, it can be accelerated by using some iterative solvers for updating \mathbf{X}'' with (28) and

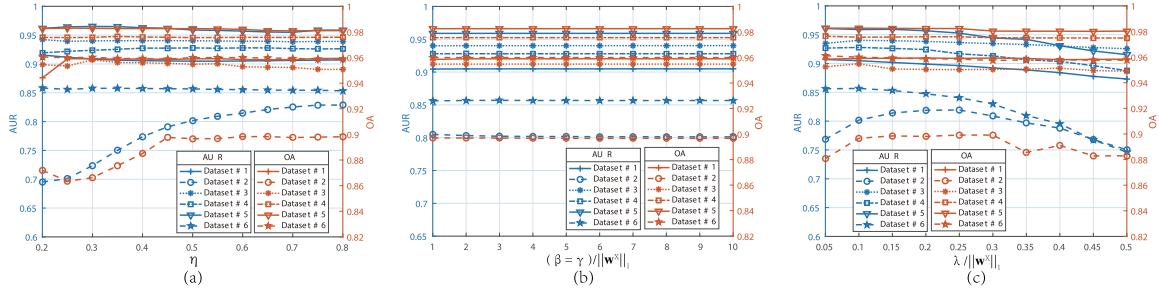


Fig. 8. Sensitivity analysis of parameters in AGSCC: (a) η , (b) β and γ , and (c) λ .

TABLE V
COMPUTATIONAL TIME (SECONDS) OF AGSCC

Data sets	N	t_{A1}	t_{A2}	t_{total}
Dataset #1 $300 \times 412 \times 1(3)$	2500	1.2	9.3	11.2
	5000	4.1	31.9	37.5
	10000	19.2	127.0	151.3
Dataset #2 $2000 \times 2000 \times 3(3)$	2500	1.1	8.7	13.0
	5000	4.3	33.8	42.7
	10000	19.4	131.2	159.5

\mathbf{Y}' with (30). The linear systems of updating \mathbf{X}'' and \mathbf{Y}' can be rewritten as

$$\begin{aligned} \mathbf{X}''^{(m)}(4w_m^X \mathbf{L}_{Y'} + 2\gamma \mathbf{I}_N) &= 2\gamma \mathbf{X}^{(m)} \\ \mathbf{Y}'^{(m)}(4w_m^{Y'}(\mathbf{L}_X + \mathbf{L}_{Y'}) + \mu \mathbf{I}_N) &= (\mu \mathbf{Y}^{(m)} + \mu \Delta^{(m)} - \mathbf{P}^{(m)}). \end{aligned} \quad (38)$$

Since the matrices of $\mathbf{L}_{Y'}$ and \mathbf{L}_X are Laplacian matrices, which are sparse, real, symmetric, and positive definite, the linear systems of (38) can be solved efficiently by using the conjugate gradient (CG) method. In addition, some preconditioners can also be used to accelerate the CG method [61], such as Jacobi, incomplete Cholesky (IC), and successive overrelaxation (SOR). In our experiments, we use the preconditioned CG with IC preconditioner for updating \mathbf{X}'' and \mathbf{Y}' .

Table V reports the computational time of AGSCC with different superpixel numbers N on Datasets #1 and #2. The algorithm is performed in MATLAB 2016a running on a Windows Laptop with Intel Core i9-10980HK CPU and 64 GB of RAM. In Table V, t_{A1} and t_{A2} represent the computational time of Algorithms 1 and 2, respectively, and t_{total} represents the total computational time of AGSCC. As can be seen from Table V, the main factor that determines the running time of AGSCC is the superpixel number N rather than the image size, and the time spent by Algorithm 2 accounts for the major part of AGSCC.

VI. CONCLUSION

In this work, we proposed an unsupervised image regression method to address the problem of heterogeneous CD. In particular, the proposed method first constructs an AG to represent the structure of preevent image, which connects each superpixel with its truly similar neighbors by using a k -selection strategy and adaptive-weighted distance metric. Based on the fact that the similarity relationship-based structure can be well

preserved across different imaging modalities, the AG can be used to translate the preevent image to the domain of postevent image with three types of regularization: forward transformation term, cycle transformation term, and sparse regularization term. By solving this structure cycle consistency-based image regression model with ADMM, a more accurate DI can be computed to measure the change level, which can be further used to calculate the CM by image segmentation methods. Extensive experiments show that the proposed method can effectively improve the detection accuracy compared with other related methods under different CD conditions.

In the future work, we will attempt to improve the computation efficiency of AGSCC and design an effective fusion strategy to fuse the forward regression (translating \mathbf{X} into \mathbf{Y}') and backward regression (translating \mathbf{Y} into \mathbf{X}') along with the cycle consistency, so as to improve the CD accuracy. Besides, in the regression model, we directly use the learned graph to complete the image regression, without exploring the high-order neighborhood information hidden in the graph. We will try to improve the regression performance by using the hypergraph and graph spectral analysis in future.

ACKNOWLEDGMENT

The authors would like to thank the editors and anonymous reviewers for their constructive suggestions that improved the presentation of this work. They also appreciate the help from Prof. Li Xinghua (Wuhan University, Wuhan, China) and Dr. Zhang Zuowei (Northwestern Polytechnical University, Xi'an, China) for sharing the codes and results.

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, 1989.
- [2] N. Longbotham *et al.*, "Multi-modal change detection, application to the detection of flooded areas: Outcome of the 2009–2010 data fusion contest," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 1, pp. 331–342, Feb. 2012.
- [3] S. Liu, D. Marinelli, L. Bruzzone, and F. Bovolo, "A review of change detection in multitemporal hyperspectral images: Current techniques, applications, and challenges," *IEEE Geosci. Remote Sens. Mag.*, vol. 7, no. 2, pp. 140–158, Jun. 2019.
- [4] 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.
- [5] Z. Lv, T. Liu, C. Shi, and J. A. Benediktsson, "Local histogram-based analysis for detecting land cover change using VHR remote sensing images," *IEEE Geosci. Remote Sens. Lett.*, vol. 18, no. 7, pp. 1284–1287, Jul. 2021.

- [6] Y. Sun, L. Lei, D. Guan, X. Li, and G. Kuang, "SAR image change detection based on nonlocal low-rank model and two-level clustering," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 293–306, 2020.
- [7] W. Ma *et al.*, "Change detection based on multi-grained Cascade forest and multi-scale fusion for SAR images," *Remote Sens.*, vol. 11, no. 2, p. 142, Jan. 2019.
- [8] Q. Wang, Z. Yuan, Q. Du, and X. Li, "GETNET: A general end-to-end 2-D CNN framework for hyperspectral image change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 1, pp. 3–13, Jan. 2018.
- [9] Z. Yuan, Q. Wang, and X. Li, "ROBUST PCANet for hyperspectral image change detection," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2018, pp. 4931–4934.
- [10] S. N. Anfinsen, L. T. Luppino, M. A. Hansen, G. Moser, and S. B. Serpico, "Unsupervised heterogeneous change detection in radar images by cross-domain affinity matching," in *Proc. IEEE Radar Conf. (RadarConf)*, Sep. 2020, pp. 1–6.
- [11] F. Bovolo, S. Marchesi, and L. Bruzzone, "A framework for automatic and unsupervised detection of multiple changes in multitemporal images," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 6, pp. 2196–2212, May 2011.
- [12] G. Moser and S. B. Serpico, "Generalized minimum-error thresholding for unsupervised change detection from SAR amplitude imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 10, pp. 2972–2982, Oct. 2006.
- [13] J. R. Jensen, E. W. Ramsey, H. E. Mackey, Jr., E. J. Christensen, and R. R. Sharitz, "Inland wetland change detection using aircraft MSS data," *Photogram. Eng. Remote Sens.*, vol. 53, no. 5, pp. 521–529, 1987.
- [14] 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.
- [15] Y. Wu, Z. Bai, Q. Miao, W. Ma, Y. Yang, and M. Gong, "A classified adversarial network for multi-spectral remote sensing image change detection," *Remote Sens.*, vol. 12, no. 13, p. 2098, Jun. 2020.
- [16] G. Mercier, G. Moser, and S. B. Serpico, "Conditional copulas for change detection in heterogeneous remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 5, pp. 1428–1441, May 2008.
- [17] J. Prendes, M. Chabert, F. Pascal, A. Giros, and J. Y. Tourneret, "A new multivariate statistical model for change detection in images acquired by homogeneous and heterogeneous sensors," *IEEE Trans. Image Process.*, vol. 24, no. 3, pp. 799–812, Mar. 2015.
- [18] M. Volpi, G. Camps-Valls, and D. Tuia, "Spectral alignment of multi-temporal cross-sensor images with automated kernel canonical correlation analysis," *J. Photogram. Remote Sens.*, vol. 107, pp. 50–63, Sep. 2015.
- [19] R. Touati and M. Mignotte, "An energy-based model encoding nonlocal pairwise pixel interactions for multisensor change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 2, pp. 1046–1058, Oct. 2018.
- [20] J. Liu, M. Gong, K. Qin, and P. Zhang, "A deep convolutional coupling network for change detection based on heterogeneous optical and radar images," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 3, pp. 545–559, Dec. 2016.
- [21] J. Liu, W. Zhang, F. Liu, and L. Xiao, "A probabilistic model based on bipartite convolutional neural network for unsupervised change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–14, 2022.
- [22] H. Li, M. Gong, M. Zhang, and Y. Wu, "Spatially self-paced convolutional networks for change detection in heterogeneous images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 4966–4979, 2021.
- [23] Y. Wu, J. Li, Y. Yuan, A. K. Qin, Q.-G. Miao, and M.-G. Gong, "Commonality autoencoder: Learning common features for change detection from heterogeneous images," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Feb. 18, 2021, doi: [10.1109/TNNLS.2021.3056238](https://doi.org/10.1109/TNNLS.2021.3056238).
- [24] X. Jiang, G. Li, X.-P. Zhang, and Y. He, "A semisupervised Siamese network for efficient change detection in heterogeneous remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–18, 2022.
- [25] Z. Liu, G. Li, G. Mercier, Y. He, and Q. Pan, "Change detection in heterogeneous remote sensing images via homogeneous pixel transformation," *IEEE Trans. Image Process.*, vol. 27, no. 4, pp. 1822–1834, Apr. 2018.
- [26] 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.
- [27] X. Li, Z. Du, Y. Huang, and Z. Tan, "A deep translation (GAN) based change detection network for optical and SAR remote sensing images," *ISPRS J. Photogramm. Remote Sens.*, vol. 179, pp. 14–34, Sep. 2021.
- [28] 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.
- [29] 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, pp. 1–22, 2022.
- [30] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 2223–2232.
- [31] Z.-G. Liu, Z.-W. Zhang, Q. Pan, and L.-B. Ning, "Unsupervised change detection from heterogeneous data based on image translation," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–13, 2022.
- [32] X. Jiang, G. Li, Y. Liu, X.-P. Zhang, and Y. He, "Change detection in heterogeneous optical and SAR remote sensing images via deep homogeneous feature fusion," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 1551–1566, 2020.
- [33] H. Dong, W. Ma, Y. Wu, J. Zhang, and L. Jiao, "Self-supervised representation learning for remote sensing image change detection based on temporal prediction," *Remote Sens.*, vol. 12, no. 11, p. 1868, Jun. 2020.
- [34] X. Li, Z. Yuan, and Q. Wang, "Unsupervised deep noise modeling for hyperspectral image change detection," *Remote Sens.*, vol. 11, no. 3, p. 258, Jan. 2019.
- [35] 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, pp. 1–21, 2022.
- [36] Y. Sun, L. Lei, X. Li, H. Sun, and G. Kuang, "Nonlocal patch similarity based heterogeneous remote sensing change detection," *Pattern Recognit.*, vol. 109, Jan. 2021, Art. no. 107598.
- [37] 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.
- [38] Y. Sun, L. Lei, X. Li, X. Tan, and G. Kuang, "Patch similarity graph matrix-based unsupervised remote sensing change detection with homogeneous and heterogeneous sensors," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 6, pp. 4841–4861, Jun. 2021.
- [39] 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, pp. 1–14, 2022.
- [40] Y. Sun, L. Lei, X. Tan, D. Guan, J. Wu, and G. Kuang, "Structured graph based image regression for unsupervised multimodal change detection," *ISPRS J. Photogramm. Remote Sens.*, vol. 185, pp. 16–31, Mar. 2022.
- [41] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, "SLIC superpixels compared to state-of-the-art superpixel methods," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 11, pp. 2274–2282, Nov. 2011.
- [42] F. Nie, X. Wang, and H. Huang, "Clustering and projected clustering with adaptive neighbors," in *Proc. 20th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2014, pp. 977–986.
- [43] L. Lei, Y. Sun, and G. Kuang, "Adaptive local structure consistency-based heterogeneous remote sensing change detection," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022.
- [44] 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.
- [45] 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, 2009.
- [46] Y. Boykov and V. Kolmogorov, "An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 9, pp. 1124–1137, Sep. 2004.
- [47] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. SMC-9, no. 1, pp. 62–66, Feb. 1979.
- [48] J. A. Hartigan and M. A. Wong, "Algorithm AS 136: A k-means clustering algorithm," *Appl. Statist.*, vol. 28, no. 1, pp. 100–108, Jan. 1979.
- [49] J. C. Bezdek, W. Full, and R. Ehrlich, "FCM: The fuzzy C-means clustering algorithm," *Comput. Geosci.*, vol. 10, nos. 2–3, pp. 191–203, 1984.
- [50] R. Touati, M. Mignotte, and M. Dahmane, "Multimodal change detection in remote sensing images using an unsupervised pixel pairwise-based Markov random field model," *IEEE Trans. Image Process.*, vol. 29, pp. 757–767, 2020.

- [51] P. Zhang, M. Gong, L. Su, J. Liu, and Z. Li, "Change detection based on deep feature representation and mapping transformation for multi-spatial-resolution remote sensing images," *ISPRS J. Photogramm. Remote Sens.*, vol. 116, pp. 24–41, Jun. 2016.
- [52] R. Touati, M. Mignotte, and M. Dahmane, "Change detection in heterogeneous remote sensing images based on an imaging modality-invariant MDS representation," in *Proc. 25th IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2018, pp. 3998–4002.
- [53] R. Touati, M. Mignotte, and M. Dahmane, "Anomaly feature learning for unsupervised change detection in heterogeneous images: A deep sparse residual model," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 588–600, 2020.
- [54] R. Touati, M. Mignotte, and M. Dahmane, "A reliable mixed-norm-based multiresolution change detector in heterogeneous remote sensing images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 9, pp. 3588–3601, Sep. 2019.
- [55] R. Touati, M. Mignotte, and M. Dahmane, "Partly uncoupled Siamese model for change detection from heterogeneous remote sensing imagery," *J. Remote Sens. GIS*, vol. 9, no. 1, pp. 1–8, 2020.
- [56] M. Yang, L. Jiao, F. Liu, B. Hou, S. Yang, and M. Jian, "DPFL-Nets: Deep pyramid feature learning networks for multiscale change detection," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, May 24, 2021, doi: [10.1109/TNNLS.2021.3079627](https://doi.org/10.1109/TNNLS.2021.3079627).
- [57] T. Zhan, M. Gong, X. Jiang, and S. Li, "Log-based transformation feature learning for change detection in heterogeneous images," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 9, pp. 1352–1356, Sep. 2018.
- [58] Y. Chen and L. Bruzzone, "Self-supervised change detection in multi-view remote sensing images," 2021, *arXiv:2103.05969*.
- [59] Z.-G. Liu, G. Mercier, J. Dezert, and Q. Pan, "Change detection in heterogeneous remote sensing images based on multidimensional evidential reasoning," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 1, pp. 168–172, Jan. 2014.
- [60] J. Wu, B. Li, W. Ni, W. Yan, and H. Zhang, "Optimal segmentation scale selection for object-based change detection in remote sensing images using Kullback–Leibler divergence," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 7, pp. 1124–1128, Jul. 2020.
- [61] F. Nar, A. Ozgur, and A. N. Saran, "Sparsity-driven change detection in multitemporal SAR images," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 7, pp. 1032–1036, Jul. 2016.



Yuli Sun received the B.S. degree from Xiangtan University, Xiangtan, China, in 2009, and the M.S. degree from the University of Science and Technology of China, Hefei, China, in 2014. He is currently pursuing the Ph.D. degree in information and communication engineering with the College of Electronic Science, National University of Defense Technology, Changsha, China.

His research interests include machine learning, remote sensing image processing, and multitemporal image analysis.



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.



and computation vision

Dongdong Guan received the B.S. degree in surveying and mapping engineering from Wuhan University, Wuhan, China, in 2013, and the M.S. degree in photogrammetry and remote sensing and the Ph.D. degree in information and communication engineering from the National University of Defense Technology, Changsha, China, in 2015 and 2019, respectively.

He works at the High-Tech Institute of Xi'an, Xi'an, China. His research interests include synthetic aperture radar image processing, machine learning, and remote sensing applications.



Junzheng Wu received the B.S. degree from Tsinghua University, Beijing, China, in 2008, and the M.S. degree from Northwestern Polytechnic University, Xi'an, China, in 2013. He is currently pursuing the Ph.D. degree in information and communication engineering with the College of Electronic Science, National University of Defense Technology, Changsha, China.

His research interests include machine learning and remote sensing image processing.



Gangyao Kuang 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 with the School of Electronic Science, National University of Defense Technology. His research interests include remote sensing, SAR image processing, change detection, synthetic aperture radar (SAR) ground moving target indication, and classification with polarimetric SAR images.



Li Liu (Senior Member, IEEE) received the Ph.D. degree in information and communication engineering from the National University of Defense Technology (NUDT), China, in 2012.

During her Ph.D. study, she spent more than two years as a Visiting Student at the University of Waterloo, Waterloo, Canada, from 2008 to 2010. From 2015 to 2016, she spent ten months visiting the Multimedia Laboratory at The Chinese University of Hong Kong. From December 2016 to November 2018, she worked as a Senior Researcher at the

Machine Vision Group, University of Oulu, Oulu, Finland. She is currently a Professor with the College of System Engineering, NUDT. Her papers have currently over 6700 citations in Google Scholar. Her current research interests include computer vision, pattern recognition, and machine learning.