

Unsupervised SAR Image Change Detection via Structure Feature-Based Self-Representation Learning

Weisong Li^{ID}, *Member, IEEE*, Yinwei Li^{ID}, *Member, IEEE*, Yiming Zhu^{ID}, *Senior Member, IEEE*, and Haipeng Wang^{ID}, *Senior Member, IEEE*

Abstract—Unsupervised low-rank matrix decomposition (LRMD) theory, leveraging inherent structural features of images, has demonstrated significant potential for synthetic aperture radar (SAR) image change detection (CD) under label scarcity. However, conventional methods rely on static sparsity priors, which inadequately model the spatial diversity and dynamic complexity of changes, leading to great decomposition errors. To address this limitation, we redefine CD within a multiview framework and propose a structurally diversified self-representation learning model. By jointly enforcing low-rank and sparse constraints on the coefficient matrix, our approach enhances the global consistency and local continuity of change information, effectively rectifying representation errors in complex change patterns. Furthermore, a robust background suppression mechanism is integrated to mitigate speckle noise and pseudo-changes, improving the discriminability between changed and unchanged regions. To refine ambiguous boundaries, an unsupervised classification refinement module is developed, calibrating transitional samples via nonlinear regression-based feature projection without labeled data. The entire framework operates without supervision, eliminating dependence on annotated samples. Extensive experiments on six bitemporal and three extended multitemporal datasets validate the effectiveness and superiority of the proposed method. The source code will be made available at <https://github.com/95xiaoli/S3CD>

Index Terms—Change detection (CD), classification refining, self-representation, structural feature, synthetic aperture radar (SAR).

Received 22 May 2025; revised 18 July 2025; accepted 9 August 2025. Date of publication 13 August 2025; date of current version 25 August 2025. This work was supported in part by the National Natural Science Foundation of China under Grant 61988102 and Grant 62271153, in part by China Postdoctoral Science Foundation (CPSF) under Grant 2025M770821, and in part by the Artificial Intelligence Promotes Research Paradigm Reform and Empowers Discipline Leap Plan of Shanghai under Grant Z-2024-302-049. (Corresponding authors: Yiming Zhu; Haipeng Wang.)

Weisong Li and Yinwei Li are with the Terahertz Technology Innovation Research Institute, Terahertz Spectrum and Imaging Technology Cooperative Innovation Center, and the School of Intelligent Emergency Management, University of Shanghai for Science and Technology, Shanghai 200093, China, and also with Shanghai Institute of Intelligent Science and Technology, Tongji University, Shanghai 200092, China (e-mail: ws_li@usst.edu.cn; liyw@usst.edu.cn).

Yiming Zhu is with the Terahertz Technology Innovation Research Institute, Terahertz Spectrum and Imaging Technology Cooperative Innovation Center, University of Shanghai for Science and Technology, Shanghai 200093, China, and also with Shanghai Institute of Intelligent Science and Technology, Tongji University, Shanghai 200092, China (e-mail: ymzhu@usst.edu.cn).

Haipeng Wang is with the Key Laboratory of Information Science of Electromagnetic Waves, Fudan University, Shanghai 200433, China (e-mail: hpwang@fudan.edu.cn).

Digital Object Identifier 10.1109/TGRS.2025.3598588

NOMENCLATURE

$\mathcal{X} \in \mathbb{R}^{d \times n \times k}$	Change feature tensor.
$\mathbf{X}_i \in \mathbb{R}^{d \times n}$	Change feature matrix of the i th view.
$P = \{P_1, P_2, \dots, P_n\}$	Set of superpixels.
$\mathbf{Z}_i \in \mathbb{R}^{n \times n}$	Coefficient matrix.
$\mathbf{z}_i$	Columns of coefficient matrix $\mathbf{Z}_i$.
$\mathbf{E}_i \in \mathbb{R}^{d \times n}$	Error matrix.
$\mathbf{A} \in \mathbb{R}^{n_1 \times d}$	Sample matrix.
$\mathbf{B} \in \mathbb{R}^{n_1 \times 2}$	Label matrix.
$\mathbf{P}$	Projection matrix.
$\mathbf{K}$	Kernel matrix.
$\ \cdot\ _0$	ℓ_0 -norm.
$\ \cdot\ _1$	ℓ_1 -norm.
$\ \cdot\ _*$	Nuclear norm.
$\ \cdot\ _F$	Frobenius norm.
$\text{tr}(\mathbf{X})$	Trace of matrix $\mathbf{X}$.

I. INTRODUCTION

RE MOTELY sensed change detection (CD) mainly focuses on identifying land cover and land use (LCLU) changes by using bitemporal or multitemporal co-registered images acquired over the same geographical area but at different times [1]. Over the past few decades, it has become one of the hottest topics in the field of remote sensing Earth observation. As an advanced active sensor, synthetic aperture radar (SAR) can operate in cloudy and rainy conditions and at any time, providing important data support for the acquisition of information on land change in harsh environments [2]. CD based on SAR imagery has been widely applied in fields such as disaster monitoring and assessment [3], [4], urban expansion monitoring [5], [6], and environmental surveillance [7], [8].

In the study of the SAR image CD problem, the analysis unit plays a crucial role in determining the minimum measurement scale in the land scene. Due to the differences in observation scenarios and the variety of resolutions of the SAR data used, the choice of an appropriate analysis granularity is essential to solve specific problems. Depending on the granularity of the analysis, SAR image CD can be broadly classified into two main categories: pixel-based methods and patch-based methods. In the evolution of SAR image CD techniques, the pixel has long been the

basic unit of processing and analysis for a long time. It is particularly suitable for low-resolution SAR data. The scattering difference characteristics of multitemporal pixel pairs have been used to detect and measure changes. After the necessary preprocessing and co-registration steps have been carried out, mathematical operators can be used directly for change comparisons, e.g., difference operators [9] and ratio operators [10].

The rapid advancement of SAR remote sensing technology has enabled abundant access to high-quality, high-resolution imagery, driving a paradigm shift in CD from pixel-based to patch-level analysis [11]. Compared to the low- and medium-resolution data, there are only a few basic scatterers in the resolution cell of the high-resolution data, so that more detailed features in the observed scene will become visible [12]. However, rich structural information also increases the spatial and temporal heterogeneity of the data, making traditional pixel-level analysis methods unsuitable for high-resolution SAR images [13]. Patch-based analysis has shown significant performance advantages in improving computational efficiency and suppressing speckle noise. Superpixels are increasingly used to better preserve spatial structure information.

Despite the great potential for the application of patch-based analysis, it also faces significant challenges. First, under high-resolution conditions, the observed scenes and targets are more refined, and the information on spatial details is richer. The contextual information of the change needs to be effectively modeled so that the change object can be characterized more comprehensive and accurately [14], [15]. Furthermore, reasonable feature representation and extraction strategies must be designed to effectively enhance the distance between the changed foreground and the unchanged background in complex change scenarios [16]. At the same time, they are able to effectively suppress the interference of pseudo-change. In a recent study, the low-rank matrix decomposition (LRMD) theory has been successfully applied to the extraction of change information [17]. The theory assumes that only a small fraction of the observed scene is changed. After decomposing the change feature matrix, the low-rank component is considered as the background part, while the sparse component is considered as the change part. The Laplacian regularization is integrated into this model to enhance the spatial correlations of the elements in the sparse matrix. In practice, the LRMD model maps the extracted change features into the same high-dimensional vector space to represent the corresponding superpixel. Information about diversity and consistency between different features has not been effectively exploited. If information from different feature spaces has large differences or inconsistencies, it will lead to serious decomposition errors.

Inspired by the data representation learning theory [18], [19], we would like to introduce the low-rank representation (LRR) theory to represent high-dimensional change feature information. As an advanced subspace learning method, LRR effectively captures the complex low-dimensional geometric structure underlying the data. Through its unique low-rank constraint mechanism, LRR profoundly reflects the intrinsic

affinity and correlation between data points (superpixels) [20]. In practice, the observations are mostly located in some low-dimensional subspace, so the data matrix stacked from the observations should have an approximate low-rank property. Based on this property, LRR and LRMD are deeply connected, both utilizing the low-rank constraint to capture the essential structure of the data. Intrinsically, LRMD uses the robust principal component analysis (RPCA) algorithm to recover the unchanged background from contaminated data perturbed by changing information [21], [22]. In contrast, our goal is to directly solve the optimization problem of the coefficient matrix, rather than recovering clean data, which directly reflects the interrelationships between the data points [23]. This facilitates the interpretation of affinities between data points and subsequent classification.

In this work, the CD scheme was formulated within the framework of multiview structure, aiming to comprehensively model spatial structural diversity [24], [25]. In general, different feature views contain low-rank structural information from different sides of the original observation data, and these unique pieces of information enhance the comprehensive description of the data [26], [27], [28]. Specifically, our framework integrates multiview features (e.g., scattering intensity, texture, and spatial structure) into a unified feature decomposition model, where a joint low-rank and sparse constraint is imposed on the coefficient matrix [29]. This constraint simultaneously ensures global consistency of homogeneous change regions and preserves local sparsity of subtle changes. Crucially, conventional low-rank models enforce static sparsity priors that cannot adapt to spatially heterogeneous change structures. This rigidity inevitably causes decomposition errors when handling clustered small-scale changes (demanding localized sparsity) versus distributed large-scale changes (requiring global consistency). This is a fundamental limitation addressed by our joint constraint paradigm.

In addition, in complex observation scenarios (e.g., densely built-up areas and seasonal rivers), where the changed region is prone to interference from speckle noise and pseudo-changes, the decomposition performance of the observation matrix inevitably degrades [17]. To address this problem, a robust background suppression technique is introduced [30], which imposes strict low-saliency constraints on the unchanged component to attenuate speckle noise and pseudo-changes. The final difference image (DI) is generated by fusing self-representation results with background-suppressed features. Furthermore, to address residual ambiguities in transitional boundaries, an unsupervised label-mapping module is proposed. This module learns nonlinear projections from high-confidence superpixel exemplars to fused difference features, thereby enabling precise value calibration for uncertain samples. The final binary change map is generated through the integration of high-confidence labeled and refined uncertain labels.

The main contributions of this article are summarized as follows.

- 1) We introduce a novel, structurally diversified self-representation learning framework, named S^3 CD,

TABLE I
DESCRIPTION OF THE NINE DATASETS USED IN THE STUDY

Dataset	Sensor	Acquisition data	Location	Polarization	Spatial resolution	Image number	Image size	Change direction
I	TerraSAR-X	10/2010,03/2011	Sendai,Japan	HH	9m×9m	2	286×298	Land→Water
II	TerraSAR-X	10/2010,03/2011	Sendai,Japan	HH	9m×9m	2	308×280	Land→Water
III	TerraSAR-X	10/2010,03/2011	Sendai,Japan	HH	9m×9m	2	941×901	Land→Water
IV	TerraSAR-X	10/2010,03/2011	Sendai,Japan	HH	9m×9m	2	1001×1001	Land→Water
V	Sentinel-1	08/2022,08/2024	Xiong'an New Area,China	VV	20m×20m	2	280×372	Land→Building
VI	Sentinel-1	08/2022,08/2024	Xiong'an New Area,China	VV	20m×20m	2	341×421	Land→Water
VII	TerraSAR-X	09/2008-03/2011	Sendai,Japan	HH	9m×9m	4	635×422	Land→Water
VIII	Sentinel-1	08/2022-08/2024	Xiong'an New Area,China	VV	20m×20m	3	491×336	Land→Water
IX	Sentinel-1	08/2022-08/2024	Xiong'an New Area,China	VV	20m×20m	3	671×462	Building→Land

specifically designed for SAR image CD. Our framework integrates multiple change feature views into a unified learning process, thereby fully exploiting and leveraging the diverse structural characteristics of the observed data to enhance the precise description of change information.

- 2) We develop a self-representation model that jointly imposes low-rank and sparse constraints to precisely measure pairwise affinities. By fusing the self-representation learning results with the background suppression results, we further improve the quality of the DI.
- 3) We construct an unsupervised classification refinement module based on nonlinear feature projections to resolve residual ambiguities in transition zones. By refining the change saliency of uncertain samples, the classification performance is further optimized.
- 4) Extensive experiments across nine diverse datasets (see Table I) demonstrate the model's effectiveness and robustness. In particular, its sustained superiority on three multitemporal datasets further validating the exceptional generalizability to long-term monitoring scenarios and practical application potential.

The remainder of this article is organized as follows. Section II briefly reviews related work on SAR image CD and LRR learning. Section III elaborates the proposed framework in detail. The extensive experimental results and detailed analysis are presented in Section IV. Finally, the conclusion is drawn in Section V. For ease of expression, we describe the main notations used in the remainder of this article as shown in Nomenclature.

II. RELATED WORK

In this section, we will briefly review the current development of CD, with a particular focus on technologies for high-resolution SAR imagery. We will also give a brief overview of LRR techniques.

A. SAR Image CD

Pixel-based CD techniques for SAR imagery are the earliest researched, the most mature and widely used. Most of the methods carry out research from the image domain to obtain change information by extracting effective change feature representations, which mainly include mathematical operators [9], [10], feature space transformation [31], [32], feature

fusion [33], [34], and deep feature extraction [35], [36]. The above methods mainly focus on bitemporal images and are limited to extracting information about changes in the spatial dimension. With the recent improvement in the availability of SAR imagery, there has been more and more research on CD in multitemporal or even long-time series SAR imagery [37]. These methods focus on how to effectively obtain rich time-dimensional information, such as the temporal extent and intensity of changes [38], [39], as well as the spatio-temporal evolution patterns [40], [41]. Strategies employed include temporal statistical measurements, change point detection, and deep time series classification.

Benefiting from the rapid development of SAR satellite remote sensing technology, high-resolution SAR images are becoming more and more abundant. Patch-based CD methods are becoming increasingly popular, which are not only able to model the contextual information of the change space more effectively, but also greatly reduce the computation [17], [42]. These approaches take the segmented superpixels or image patches (such as square patches) as the basic unit of analysis and assign binary change labels to them, which are the same for all pixels within the same object block [43], [44]. Recently, low-rank modeling theory and saliency detection techniques have been incorporated into superpixel analysis methods and have achieved certain performance advantages [45], [46], [47]. In [17], the RPCA algorithm with normal graph Laplacian regularization constraints was proposed, and a further classification refinement module was constructed for the detection of boundary superpixels refinement. The model operates entirely under unsupervised conditions and shows a certain degree of robustness. However, existing models rely on idealized sparsity assumptions for changes, overlooking the complexity and dynamic diversity of change spaces, which results in significant performance degradation when change regions violate approximate sparsity conditions. In contrast, we incorporate the self-representation model into a multiview framework through the joint enforcement of low-rank and sparse constraints. This explicitly models heterogeneous change structures while ensuring global consistency and local sensitivity, eliminating LRMD's performance degradation under nonsparse change patterns.

B. Low-Rank Representation

In this section, we will briefly review the basic model of LRR. It assumes that the observed data is sampled from a superposition or concatenation of several low-dimensional

subspaces [18]. Consider a set of d -dimensional data samples $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n] \in \mathbb{R}^{d \times n}$ drawn from a union of multiple subspaces $\{\mathcal{S}_i\}_{i=1}^k$. The goal of LRR is to jointly find the lowest-rank representation of the $\mathbf{Y}$, which is defined as the following rank minimization problem:

$$\min_{\mathbf{Z}, \mathbf{E}} \text{rank}(\mathbf{Z}) + \lambda \|\mathbf{E}\|_0 \quad \text{s.t. } \mathbf{X} = \mathbf{DZ} + \mathbf{E} \quad (1)$$

where $\mathbf{D} = [\mathbf{D}_1, \mathbf{D}_2, \dots, \mathbf{D}_n] \in \mathbb{R}^{d \times n}$ is a given bases or dictionary used to linearly represent the original data, $\mathbf{E}$ denotes the error term representing the noise, $\|\cdot\|_0$ is the ℓ_0 -norm, and $\lambda > 0$ is a tradeoff parameter for balancing the low rank term and the error term.

Since it is a nonconvex and NP-hard problem, the alternative convex relaxation to solve that is

$$\min_{\mathbf{Z}, \mathbf{E}} \|\mathbf{Z}\|_* + \lambda \|\mathbf{E}\|_1 \quad \text{s.t. } \mathbf{X} = \mathbf{DZ} + \mathbf{E} \quad (2)$$

where $\|\cdot\|_*$ is the nuclear norm and $\|\cdot\|_1$ denotes the ℓ_1 -norm to promote sparsity.

When we directly solve for the clean matrix $\mathbf{L} = \mathbf{DZ}$ of the original data, instead of the coefficient matrix $\mathbf{Z}$, (2) is transformed into an optimization expression for RPCA [21], [22]

$$\min_{\mathbf{L}, \mathbf{E}} \text{rank}(\mathbf{L}) + \lambda \|\mathbf{E}\|_1 \quad \text{s.t. } \mathbf{X} = \mathbf{L} + \mathbf{E}. \quad (3)$$

It can be seen that the purpose of RPCA is to recover the low-rank clean part from the original contaminated data, which implicitly assumes that the main structures in the original data are in a single low-rank subspace. Unlike this, LRR takes multiple subspaces into account by introducing a dictionary matrix $\mathbf{D}$ on top of RPCA to achieve a better decomposition. In many cases, the original data matrix $\mathbf{X}$ is chosen as $\mathbf{D}$ to better explain the intrinsic affinity between data points, then (2) can be expressed as

$$\min_{\mathbf{Z}, \mathbf{E}} \|\mathbf{Z}\|_* + \lambda \|\mathbf{E}\|_1 \quad \text{s.t. } \mathbf{X} = \mathbf{XZ} + \mathbf{E}. \quad (4)$$

III. METHODOLOGY

In this section, we will describe the $S^3\text{CD}$ algorithm and its specific implementation process. First, the input bi-temporal or multitemporal SAR images are preprocessed, and the change feature tensor is then obtained by superpixel segmentation. Second, we construct a self-representation model with joint low-rank and sparse constraints to extract the coefficient matrix $\mathbf{Z}$. The model can be optimized using the linearized alternating direction method with adaptive penalty (LADMAP) algorithm [48]. Furthermore, a robust background suppression algorithm is introduced to fuse the decomposition results. In addition, a label mapping module is designed to refine the change saliency of uncertain samples through nonlinear regression-based feature projection. Fig. 1 shows the overall framework of the proposed method. Note that our approach is completely unsupervised and does not require any additional real labels or supervisory information, which completely distinguishes it from current popular deep learning algorithms.

A. Data Preprocessing and Feature Extraction

In the first stage, the input SAR image data undergo comprehensive preprocessing to mitigate interference from multitemporal imaging differences and enhance genuine change signals. This includes radiometric calibration, topographic correction, and precision co-registration. Critically, effective speckle noise suppression is indispensable for reliable CD. Following established practices [41], we apply multitemporal filtering, which jointly leveraging spatial and temporal correlations, to suppress coherent noise while preserving structural details. Given that preprocessing constitutes a standard procedure rather than a core contribution, technical specifics are omitted here.

The next important step is image segmentation. We choose superpixels instead of regular blocks to preserve the contextual information effectively. The computational complexity of matrix decomposition is greatly reduced after superpixel segmentation. Same as the previous work [17], the classical simple linear iterative clustering (SLIC) algorithm is used to generate multiple irregular superpixels. Considering that the extensive speckle noise present in SAR images can seriously degrade the segmentation accuracy, SLIC acts on the initial DI maps obtained by the LR operator (for bi-temporal images) or the cumulative LR operator (for multitemporal images) rather than on the original image.

Once the segmented set of superpixels $P = \{P_1, P_2, \dots, P_n\}$ is obtained, the change feature tensor can be constructed. We systematically construct the change feature tensor through three complementary feature extraction approaches. First, we utilize the fundamental scattering intensity difference as it directly captures the backscattering variations caused by land surface property changes, serving as the most intuitive change indicator. Second, we employ steerable pyramid filters with a three-layer, four-direction configuration (totaling 12 dimensions) to capture multiscale directional texture patterns. Third, we implement Gabor filters with three scales and 12 directions (yielding 36-dimensional features) to comprehensively characterize structural changes at different orientations and resolutions. These carefully selected filter banks work synergistically, together forming a robust multidimensional feature representation that effectively preserves both global and local change characteristics in the DI.

For each feature space, the extracted features are stacked vertically, then we can construct the change feature tensor $\mathcal{X} = [\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_k] \in \mathbb{R}^{d \times n \times k}$, where $\mathbf{X}_k$ is the change feature matrix in k th feature space, d is the total dimension of the feature, n is the number of the superpixels, and k is the number of the feature space.

B. Low-Rank Self-Representation Modeling

For the given change feature tensor $\mathcal{X}$, each feature subspace contains feature information describing a different aspect of the change in the observed region. This information from different subspaces has both features that belong to its own independent approximate low-dimensional structure associated with the change, and contains common consistency information. When the percentage of change in the observed region is

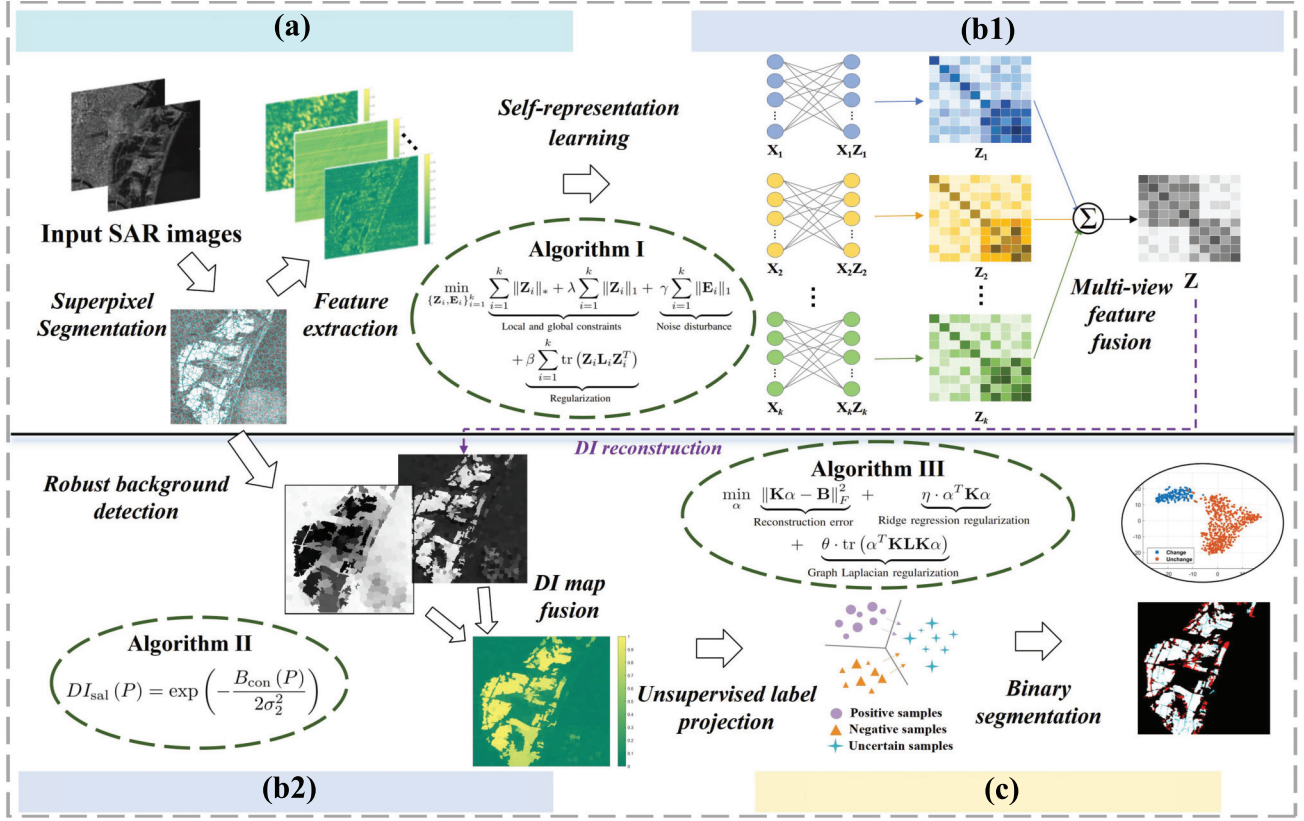


Fig. 1. Framework of the proposed S³CD algorithm. It can be divided into three main parts. Superpixel segmentation and feature extraction are first performed. Then, the self-representation model is constructed to obtain the coefficient matrix $\mathbf{Z}$, and a robust background detection module is introduced to improve the separability between changed and unchanged areas. Finally, an unsupervised label projection module is designed to refine the fused DIs and generate the binary change maps. (a) Preprocessing and feature extraction. (b1) Low-rank self-representation modeling. (b2) Background saliency detection. (c) Classification refinement.

unknown, the feature matrix $\mathbf{X}$ for each view can be decomposed into two components. The first component represents low-rank structural information (containing both background and changed regions), while the second corresponds to sparse noise perturbation. Then, the generic optimization problem can be expressed as

$$\min_{\{\mathbf{Z}_i, \mathbf{E}_i\}_{i=1}^k} \sum_{i=1}^k f(\mathbf{X}_i, \mathbf{Z}_i) + \gamma \sum_{i=1}^k g(\mathbf{X}_i, \mathbf{E}_i) + \beta \sum_{i=1}^k \Theta(f, g) \quad (5)$$

where f is a function of obtaining the low-rank component from the original data, and g is used to extract the perturbation term. Θ is the introduced regularization term to make it easier to separate both two components. The parameters γ and β are used to balance the sparsity and regularization term.

Diversity and consistency between different feature subspaces are important for improving the performance of LRR. Diversity implies that different feature views provide rich complementary information from different perspectives. In (5), the self-representation strategy can effectively preserve the diversity information. Consistency, on the other hand, reflects the inherent consensus among multiple feature subspaces, which ensures that different feature views share a uniform and stable structure in representing the original data. Given the effectiveness of the low-rank structure in recovering the low-dimensional structure of the underlying data, the coefficient

matrix $\mathbf{Z}$ of each feature view exhibits an approximate block-diagonal structure, which ensures consistency across views. In addition, the introduction of the regularization term ensures that the results do not deviate excessively to one side, it also allows for appropriate local smoothing of the results and reduces the problem of the hole and gap effect.

In the concrete implementation of (5), we first impose both low-rank and sparsity constraints on the coefficient matrix $\mathbf{Z}$. Without considering the regularization term, (5) can be rewritten as

$$\min_{\{\mathbf{Z}_i, \mathbf{E}_i\}_{i=1}^k} \sum_{i=1}^k \|\mathbf{Z}_i\|_* + \lambda \sum_{i=1}^k \|\mathbf{Z}_i\|_1 + \gamma \sum_{i=1}^k \|\mathbf{E}_i\|_1$$

$$\text{s.t. } \mathbf{X}_i = \mathbf{X}_i \mathbf{Z}_i + \mathbf{E}_i, \quad \mathbf{Z}_i \geq 0. \quad (6)$$

The low-rank criterion focuses on capturing the overall or global structure of the original data, while the sparsity criterion focuses more on capturing the local or detailed structure. Applying both constraints to $\mathbf{Z}$ can more effectively fuse the global and local structure information, thus capturing the complexity of the data comprehensively and accurately. Therefore, the optimal representation of the model exhibits a structural close to the block diagonal. Fig. 2 provides three examples of a visual comparison of the affinity graphs obtained by the proposed model with those obtained by (3). One can clearly observe that the block-diagonal structure

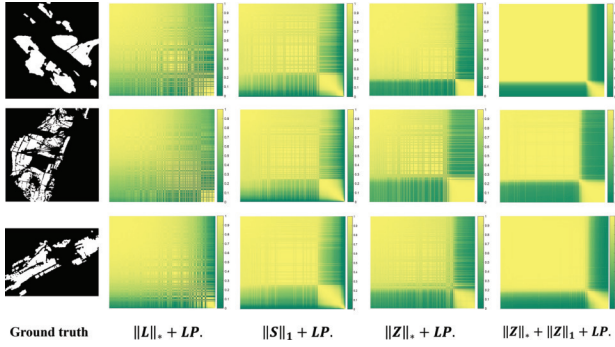


Fig. 2. Comparisons of affinity graphs generated by different terms. From top to bottom, they correspond to datasets I, VI, and VII, respectively. From left to right, they correspond to ground truth and the generated affinity graphs by different terms, where $\mathbf{L}$ and $\mathbf{S}$ are the low-rank and sparsity terms obtained from (3), respectively, and LP means using Laplacian regularization.

of the affinity matrix obtained by (3) is rather blurred. In contrast, the proposed method can provide a relatively distinct block-diagonal structure with little noise, which is helpful for subsequent data classification.

As for the regularization term in (5), based on graph theory, most studies use the Laplacian regularization term, which induces adjacent or similar data points to converge in the representation coefficients, thus effectively enhancing the similarity in the local space. The spatial similarity can be represented by the similarity matrix $\mathbf{W}$. Considering that the columns of $\mathbf{Z}$ are the self-representations of the data points with the original data $\mathbf{X}$ as the dictionary, thus the regularization term can be defined as

$$\Theta(f, g) = \sum_{i,j=1}^N \|\mathbf{z}_i - \mathbf{z}_j\|_2^2 w_{i,j} \quad (7)$$

where w denotes the weight that connects the coefficient vector $\mathbf{z}_i$ and $\mathbf{z}_j$. In more cases, the complex relationships between data points cannot be fully described by pairwise relationships of the normal graph. Therefore, we replace the original normal Laplacian regularizer with a hyper-Laplacian regularizer [49]. Based on hypergraph theory, the normal weight $\mathbf{W}$ is then weighted by the sum of the distances d_e between data points within each hyperedge $\mathbf{e}$. (7) can be rewritten as

$$\begin{aligned} \Theta(f, g) &= \sum_{i,j=1}^N \|\mathbf{z}_i - \mathbf{z}_j\|_2^2 \frac{w_e}{d_e} \\ &= \text{tr}(\mathbf{ZDZ}^T) - \text{tr}(\mathbf{ZWZ}^T) = \text{tr}(\mathbf{ZLZ}^T) \end{aligned} \quad (8)$$

where $\text{tr}(\cdot)$ denotes the trace of a matrix, $\mathbf{D}$ is a degree matrix, and $\mathbf{L}$ is the Laplacian matrix built on $\mathbf{W}$.

By integrating the hyper-Laplacian regularization term, the final tensorial low-rank self-representation model is formulated as

$$\min_{\{\mathbf{Z}_i, \mathbf{E}_i\}_{i=1}^k} \underbrace{\sum_{i=1}^k \|\mathbf{Z}_i\|_* + \lambda \sum_{i=1}^k \|\mathbf{Z}_i\|_1}_{\text{Local and global constraints}} + \underbrace{\gamma \sum_{i=1}^k \|\mathbf{E}_i\|_1}_{\text{Noise disturbance}}$$

$$\begin{aligned} &+ \underbrace{\beta \sum_{i=1}^k \text{tr}(\mathbf{Z}_i \mathbf{L}_i \mathbf{Z}_i^T)}_{\text{Regularization}} \\ \text{s.t. } &\mathbf{X}_i = \mathbf{X}_i \mathbf{Z}_i + \mathbf{E}_i, \quad \mathbf{Z}_i \geq 0 \end{aligned} \quad (9)$$

where λ , γ , and β are the balancing parameters. In this model, the local and global structural information is effectively fused by simultaneously imposing the dual constraints of low rank and sparsity on the coefficient matrix $\mathbf{Z}$. This not only significantly improves the ability to detect continuously changing regions, but also greatly enhances its sensitivity to subtle, locally changing regions. In addition, the hyper-Laplacian regularization term is an extension of the normal Laplacian regularization, which is able to capture the similarity information between the data points more comprehensively, and also helps the model to remain robust when dealing with a highly noisy image. For the nonnegativity constraint, i.e., $\mathbf{Z} \geq 0$, it means that when we use $\mathbf{Z}$ to represent the original data $\mathbf{X}$, each data point is composed of a nonnegative linear combination of other data points plus sparse errors $\mathbf{E}$. This constraint offers two key advantages: First, the nonnegative combination yields more interpretable and intuitive model outputs. Second, it inherently enhances sparsity, as certain elements are constrained to zero, the sparsity of $\mathbf{Z}$ naturally increases.

Optimization: To solve the optimization problem of (9), the LADMAP is adopted. Since the objective function of (9) is inseparable, the auxiliary variable $\mathbf{J}$ is introduced, and then the optimization problem of (9) can be transformed into an equivalent problem as follows:

$$\begin{aligned} \min_{\mathbf{Z}_i, \mathbf{E}_i} & \|\mathbf{Z}_i\|_* + \lambda \|\mathbf{J}_i\|_1 + \gamma \|\mathbf{E}_i\|_1 + \beta \text{tr}(\mathbf{Z}_i \mathbf{L}_i \mathbf{Z}_i^T) \\ \text{s.t. } & \mathbf{Y}_i = \mathbf{Y}_i \mathbf{Z}_i + \mathbf{E}_i, \quad \mathbf{Z}_i = \mathbf{J}_i, \quad \mathbf{Z}_i \geq 0. \end{aligned} \quad (10)$$

We rewrite (10) as minimizing the following augmented Lagrangian function:

$$\begin{aligned} \mathcal{L}(\mathbf{Z}_i, \mathbf{J}_i, \mathbf{E}_i, \mathbf{Y}_i^1, \mathbf{Y}_i^2, \mu) &= \|\mathbf{Z}_i\|_* + \lambda \|\mathbf{J}_i\|_1 + \gamma \|\mathbf{E}_i\|_1 \\ &+ \beta \text{tr}(\mathbf{Z}_i \mathbf{L}_i \mathbf{Z}_i^T) \\ &+ \text{tr} \left[(\mathbf{Y}_i^1)^T (\mathbf{X}_i - \mathbf{X}_i \mathbf{Z}_i - \mathbf{E}_i) \right] + \text{tr} \left[(\mathbf{Y}_i^1)^T (\mathbf{Z}_i - \mathbf{J}_i) \right] \\ &+ \frac{\mu}{2} (\|\mathbf{X}_i - \mathbf{X}_i \mathbf{Z}_i - \mathbf{E}_i\|_F^2 + \|\mathbf{Z}_i - \mathbf{J}_i\|_F^2) \end{aligned} \quad (11)$$

where $\mathbf{Y}_i^1$ and $\mathbf{Y}_i^2$ are the Lagrange multipliers, and μ is the penalty parameter. The variables $\mathbf{Z}_i$, $\mathbf{J}_i$, $\mathbf{E}_i$, $\mathbf{Y}_i^1$, and $\mathbf{Y}_i^2$ can be optimized iteratively by minimizing the augmented Lagrangian function $\mathcal{L}$, and in each iteration, the variables are updated alternately. The detailed derivation of LADMAP is given in the Appendix. The main steps of solving the optimization problem in (9) are summarized in Algorithm 1.

After decomposition for the original data, the novel DI map can be constructed from the resulting coefficient matrix set $\mathbf{Z} = [\mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}_k] \in \mathbb{R}^{n \times n \times k}$. Specifically, the difference values of the corresponding superpixels are evaluated using the cumulative sum of the ℓ_1 -norm of the each column of $\mathbf{Z}$: $\sum_{i=1}^k \|\mathbf{z}_i\|_1$. The larger of the $\sum_{i=1}^k \|\mathbf{z}_i\|_1$, the higher the value

Algorithm 1 Solving Problem (9) by LADMAP

Input: Data tensor $\mathcal{X} = \{\mathbf{X}_i\}_{i=1}^k$, parameters λ, γ , and β

- 1 Initialize $\{\mathbf{Z}_i = \mathbf{J}_i = \mathbf{E}_i = 0\}_{i=1}^k, \{\mathbf{Y}_i^1 = \mathbf{Y}_i^2 = 0\}_{i=1}^k$,
 $\lambda = 0.5, \gamma = 1, \beta = 1, \mu^0 = 10^{-6}, \mu_{\max} = 10^6$,
 $\rho_0 = 2.5, \varepsilon_1 = 10^{-6}, \varepsilon_2 = 10^{-2}$;
- 2 **while** not converged ($t = 0, 1, \dots$) **do**
- 3 **for** $i = 1: k$ **do**
- 4 Fix others and update $\mathbf{Z}_i$ by solving (26)
- 5 Fix others and update $\mathbf{J}_i$ by solving (27)
- 6 Fix others and update $\mathbf{E}_i$ by solving (28)
- 7 Fix others and update $\mathbf{Y}_i^1$ and $\mathbf{Y}_i^2$ by
solving (29) and (30), respectively.
- 8 **end**
- 9 Update μ : $\mu_{t+1} = \min(\mu_{\max}, \rho_k \mu_k)$
- 10 Check the convergence conditions:
- 11 **for** $i = 1: k$ **do**
- 12 **if** $\|\mathbf{X}_i - \mathbf{X}_i \mathbf{Z}_i - \mathbf{E}_i\| / \|\mathbf{X}_i\| < \varepsilon_1$ and
 $\max\{\mu \eta \|\mathbf{Z}_i^{t+1} - \mathbf{Z}_i^t\|, \mu \|\mathbf{J}_i^{t+1} - \mathbf{J}_i^t\|, \mu \|\mathbf{E}_i^{t+1} - \mathbf{E}_i^t\|\} < \varepsilon_2$ **then**
- 13 **break**;
- 14 **end**
- 15 **end**
- 16 **end**

Output: $\mathbf{Z}^{t+1}$;

assigned to the superpixel. After normalization, a novel gray-scale DI map is generated.

C. Background Saliency Detection

Previous approaches predominantly focus on foreground difference enhancement through saliency maximization while ignoring the important role of background structure information. Instead, we consider it from the perspective of suppressing the unchanged information in the background, and implement robust background detection techniques to accurately isolate those regions with small change differences. This not only highlights the real change information, but also greatly broadens the decision boundary between the changed and unchanged regions.

Our approach is based on the work of boundary connectivity [30]. An undirected weighted graph is first constructed by connecting all neighboring superpixels (P_i, P_j) and defining the edge weights $d_{\text{euc}}(P_i, P_j)$ as the Euclidean distance from the mean scattering amplitude of all the pixels within the superpixel. Then, the geodesic distance between any two superpixels is defined as

$$d_{\text{geo}}(P_i, P_j) = \min_{P_1, P_2, \dots, P_N} \sum_{i=1}^{N-1} d_{\text{euc}}(P_i, P_{i+1}) \quad (12)$$

where N denotes the number of superpixels. The above formula shows that if we find the shortest path between P_i and P_j , then d_{geo} is the sum of the weights of all edges on this

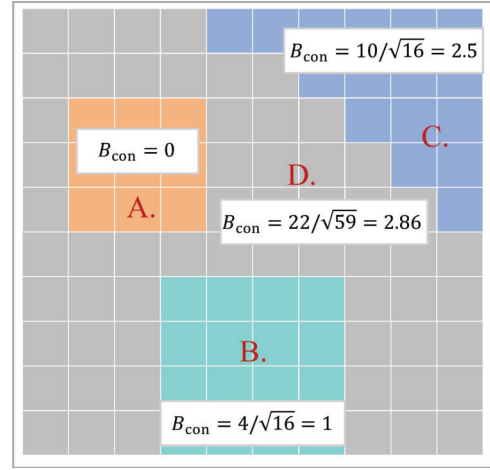


Fig. 3. Illustrative example of the boundary connectivity in different regions. Areas A and B denote the real changed regions, area C denotes the pseudo-changed region, and area D represents the background region. The values of boundary connectivity are calculated by (15).

path. Then, the area similarity and boundary similarity can be defined, respectively,

$$S_{\text{area}}(P) = \sum_{i=1}^N \exp\left(-\frac{d_{\text{geo}}^2(P_i, P_j)}{2\sigma_1^2}\right) \quad (13)$$

$$S_{\text{bnd}}(P) = \sum_{i=1}^N \exp\left(-\frac{d_{\text{geo}}^2(P, P_i)}{2\sigma_1^2}\right) \cdot \delta(P_i \in \mathbb{U}_{\text{bnd}}) \quad (14)$$

where σ_1 is an adjusting parameter and $\delta(\cdot)$ is set to 1 when the superpixels on the boundary and 0 for otherwise. Based on (13) and (14), the boundary connectivity can be represented as

$$B_{\text{con}}(P) = \frac{S_{\text{bnd}}(P)}{\sqrt{S_{\text{area}}(P)}}. \quad (15)$$

In the analysis of spatial layout, the changed regions tend to be sparsely connected to image boundaries. In contrast, stable background regions tend to form tighter and more coherent connections with image boundaries. Based on this observation, we can effectively quantify the probability that a region is attributed to the background by evaluating the strength of boundary similarity, as elaborated in (15). Specifically, if a given superpixel shows a high degree of boundary connectivity, i.e., its boundary is strongly and uniformly associated with the surrounding pixels, it is highly likely that the superpixel corresponds to a flat background region in the image. This phenomenon shows high reliability in changed scenes. Fig. 3 shows an illustrative example of boundary connectivity. For the real changed regions A and B, they have low values of boundary connectivity, while for the pseudo-changed region C and background region D, they have high values. Given the constraints of the unsupervised learning condition, comparing the connectivity of the superpixel to be evaluated with the boundary superpixel becomes an efficient and practical strategy to measure the background saliency of a region. In short, if a superpixel has strong boundary connectivity characteristics, it is more likely to be judged as a background part; conversely, if the boundary connectivity is weak, it is more likely to be a

changed region. Therefore, the change saliency can be further expressed as

$$DI_{\text{sal}}(P) = \exp\left(-\frac{B_{\text{con}}(P)}{2\sigma_2^2}\right). \quad (16)$$

Upon obtaining the change saliency map DI_{sal} using robust background detection, we propose a dynamic weight allocation mechanism for decision-level fusion. This adaptive fusion strategy synergistically combines the structural coherence preserved in the LRR-derived DI map DI_{base} with the salient foreground features highlighted in DI_{sal} . The fusion process is formulated as

$$DI_{\text{fuse}} = \mu DI_{\text{base}} + (1 - \mu) DI_{\text{sal}} \quad (17)$$

where $\mu \in [0, 1]$ is a balanced parameter to assign different weights to two terms. Such adaptive decision-level fusion strategically leverages complementary strengths of two specialized processes: The self-representation model [see (9)] operates within a feature-driven paradigm, learning latent change structures through multiview tensor decomposition regularized by low-rank and sparse constraints. In contrast, the background suppression module functions in the spatial-topological domain, leveraging boundary connectivity metrics to quantify background stability via geodesic distances between superpixels. This preserves computational efficiency and further ensures operational flexibility, allowing real-time weight adjustment μ across diverse scenarios: for large-scale changed regions (extensive spatial distribution), we enforce $\mu \rightarrow 1$ to emphasize structural consistency preservation through low-rank constraints; conversely, for small-proportion changes (localized distribution), we prioritize $\mu \rightarrow 0$ to amplify background suppression capability against pseudo-changes.

D. Unsupervised Classification Refinement

After obtaining the fused superpixel DI map, residual segmentation ambiguity persists, particularly in transitional boundary regions. To address this challenge, we develop an unsupervised classification refinement module based on label-projection learning. This module establishes a nonlinear mapping from the original feature tensor to fused difference values by leveraging high-confidence superpixel exemplars, thereby enabling precise value calibration for ambiguous samples. The technical implementation of this refinement module will be elaborated in this section. Specifically, we propose a kernelized graph-regularized regression framework. The refinement process is formulated as an unsupervised regression problem, where only high-confidence positive and negative samples are utilized for model training, while uncertain samples are inferred through the learned nonlinear mapping and graph-based regularization.

The fused DI map, enhanced by spatial structural constraints and background irrelevant information suppression, ensures high confidence levels in the majority of its superpixel samples. We classify superpixels into three categories using two thresholds τ_1 and τ_2 ($\tau_1 < \tau_2$): positive samples (changed), negative samples (unchanged) and uncertain samples, as shown in Fig. 4. Then, the high-confident sample

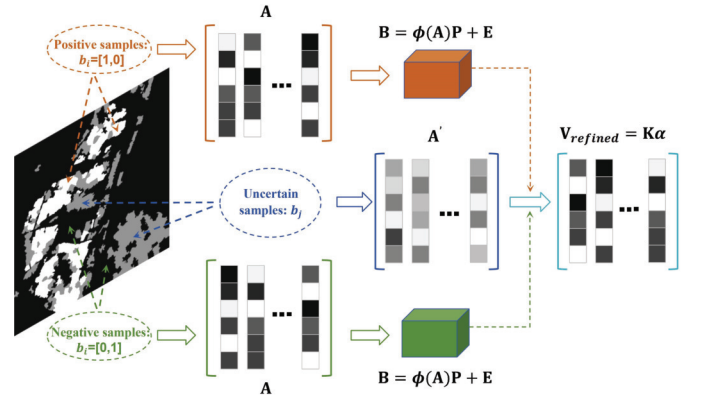


Fig. 4. Illustration of the classification refinement. All the positive samples (in white), negative samples (in black), and uncertain samples (in gray) are selected from the fused DI. Then, we refine the change saliency of those uncertain samples based on unsupervised label projection learning.

change feature matrix $\mathbf{A} = [a_1, a_2, \dots, a_{n_1}]^T \in \mathbb{R}^{n_1 \times d}$ can be constructed, where n_1 is the number of selected positive and negative samples. Correspondingly, the binary label matrix $\mathbf{B} = [b_1, b_2, \dots, b_{n_1}] \in \mathbb{R}^{n_1 \times 2}$ encodes sample categories: $b_i = [1, 0]$ for positives and $[0, 1]$ for negatives. Uncertain samples $\mathbf{A}' \in \mathbb{R}^{n_2 \times d}$ are reserved for subsequent refinement.

To keep nonlinear separability and enforce spatial consistency, we define the projection learning as a kernel ridge regression problem regularized by a graph Laplacian term. Let $\phi: \mathbb{R}^d \rightarrow \mathcal{H}$ map row features to a reproducing kernel Hilbert space. The projection model becomes

$$\mathbf{B} = \phi(\mathbf{A})\mathbf{P} + \mathbf{E} \quad (18)$$

where $\mathbf{P} \in \mathcal{H}^2$ is a projection matrix and $\mathbf{E}$ is the error term. Using the representer theorem, $\mathbf{P}$ can be further expressed as $\mathbf{P} = \phi(\mathbf{A})^T \alpha$, where α is the dual coefficient matrix. Then, we construct a graph $\mathcal{G}$ over all superpixel (including uncertain ones) to encode spatial adjacency and feature similarity. The affinity matrix $\mathbf{W}$ can be defined as: $w_{ij} = \exp(-\|a_i - a_j\|^2 / 2\sigma^2)$, where σ is the bandwidth parameter. The graph Laplacian matrix is $\mathbf{L} = \mathbf{D} - \mathbf{W}$, where $\mathbf{D}$ is the degree matrix. The graph Laplacian regularization term, which enforces smoothness of the projected features, can be expressed as

$$\Theta_{\text{graph}} = \text{tr}(\alpha^T \mathbf{K} \mathbf{L} \mathbf{K} \alpha) \quad (19)$$

where $\mathbf{K} = \phi(\mathbf{A})\phi(\mathbf{A})^T$ is the kernel matrix.

Based on the ridge regression framework, by intergrating the graph Laplacian regularization term, the final unified objective function is constructed as

$$\begin{aligned} \min_{\alpha} \quad & \underbrace{\|\mathbf{K}\alpha - \mathbf{B}\|_F^2}_{\text{Reconstruction error}} + \underbrace{\eta \cdot \alpha^T \mathbf{K} \alpha}_{\text{Ridge regression regularization}} \\ & + \underbrace{\theta \cdot \text{tr}(\alpha^T \mathbf{K} \mathbf{L} \mathbf{K} \alpha)}_{\text{Graph Laplacian regularization}} \end{aligned} \quad (20)$$

where η and θ are two balance parameters. The model integrates three critical components to achieve robust refinement in difference value estimation. The first term minimizes the reconstruction error, ensuring fidelity to reliable training data

Algorithm 2 The Proposed S³CD Algorithm

Input: Change feature tensor $\mathcal{X} = \{\mathbf{X}_i\}_{i=1}^k$, parameters $\lambda, \gamma, \beta, \mu, \theta$ and η .

- 17 Solve problem (9) by Algorithm 1, and obtain the optimal solution $\mathbf{Z} = \{\mathbf{Z}_i\}_{i=1}^k$.
- 18 Calculate the $\mathbf{Z}$: $\sum_{i=1}^k \|\mathbf{z}_i\|_1$ and obtain the DI_{base} .
- 19 Calculate the change saliency by (16), and obtain the DI_{sal} .
- 20 Fuse the above two DI maps by (17), and generate the DI_{fuse} .
- 21 Solve problem (21) and generate the DI_{RF} .
- 22 Classify the DI_{RF} using (22).

Output: The binary classification results $\mathbf{C}_{\text{final}}$.

while preserving discriminative patterns. The second term introduces a nonlinear ridge regularization on the kernel-space coefficients, effectively controlling model complexity by penalizing excessively large parameter values and mitigating overfitting risks. The third term incorporates spatial smoothness through the graph Laplacian matrix $\mathbf{L}$. This matrix enforces local consistency by ensuring that adjacent superpixels with similar features maintain proximate difference values in the projected subspace, thereby suppressing fragmented artifacts and enhancing regional homogeneity. The balancing parameters η and θ orchestrate the tradeoff among regularization strength, spatial coherence, and label fidelity, collectively optimizing the model for accurate and consistent label projection in complex SAR scenarios.

Differentiating the objective function (20) and we can get the following closed-form solution:

$$\alpha = (\mathbf{K} + \eta \mathbf{I} + \theta \mathbf{L}\mathbf{K})^{-1} \mathbf{B}. \quad (21)$$

The learned projection is then applied to uncertain samples $\mathbf{A}'$, and the refined difference values $\mathbf{V}_{\text{refined}} = \mathbf{K}\alpha \in \mathbb{R}^{(n-n_1) \times 2}$ can be obtained. Then, the binary labels for uncertain samples are then assigned as

$$b'_j = \begin{cases} 1 \text{ (changed)}, & \text{if } b'_j(1) \geq b'_j(2) \\ 0 \text{ (unchanged)}, & \text{otherwise.} \end{cases} \quad (22)$$

It is critical to note that the condition $b'_i(1) \geq b'_i(2)$ (using $\geq$ instead of $>$) explicitly includes equality cases as changes to prevent missed detections, thereby mitigating missed detection at the expense of tolerating higher false alarms, as practical scenarios prioritize minimizing omissions over false alerts.

Let $\mathbf{C}_{\text{high}} = \{(c_i, b_i)\}_{i=1}^{n_1}$ represent high-confident samples, where c_i is the spatial index. The uncertain samples' labels $\mathbf{C}_{\text{uncertain}} = \{(c_j, b_j)\}_{j=1}^{n-n_1}$ are merged with $\mathbf{C}_{\text{high}}$ to construct the full binary change map $\mathbf{C}_{\text{final}}$

$$\mathbf{C}_{\text{final}}(x) = \begin{cases} b_i, & \text{if } x = c_i \in \mathbf{C}_{\text{high}} \\ b'_j, & \text{if } x = c_j \in \mathbf{C}_{\text{uncertain}}. \end{cases} \quad (23)$$

The complete superpixel CD process of S³CD is summarized in Algorithm 2.

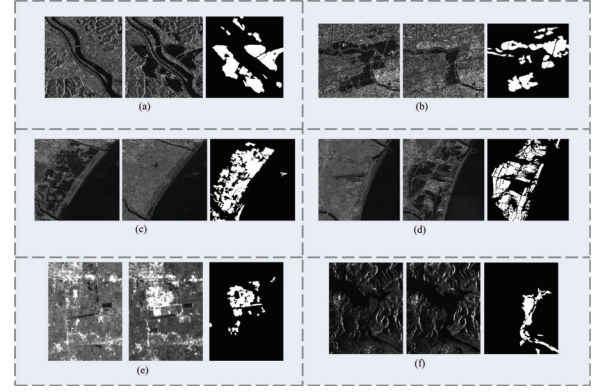


Fig. 5. Bitemporal images and corresponding reference images. From (a) to (f), they correspond to datasets I–VI, respectively.

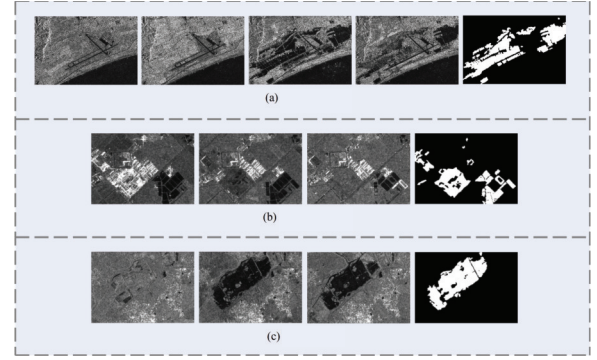


Fig. 6. Multitemporal images and corresponding reference images. From (a) to (c), they correspond to datasets VII–IX, respectively.

IV. EXPERIMENTS

To comprehensively evaluate the performance of the proposed method, extensive experiments are conducted on nine real-world datasets. The proposed method is first compared with some state-of-the-art (SOTA) CD methods to verify its superiority. The effectiveness of the algorithmic architecture is then demonstrated through ablation studies. Finally, the parameter sensitivity and computational complexity of the algorithm are discussed.

A. Dataset and Experimental Setting

1) *Datasets*: The nine datasets used in the study are from the TerraSAR-X and Sentinel-1 satellite sensors, respectively. The first six of them are bitemporal, as shown in Fig. 5, and the last three are multitemporal, as shown in Fig. 6. The bitemporal datasets can comprehensively evaluate the detection performance of the model, while the multitemporal datasets are used to test the generalization ability and scalability. The case areas are located in Sendai, Japan, and Xiong'an New Area, China. The changes in the former are mainly due to sudden tsunami damage, while in the latter are mainly belonged to seasonal land cover disturbance. Table I provides details of these datasets, revealing their varied characteristics. Notably, they provide different spatial resolutions, including 9 and 20 m, image sizes ranging from 200 to 1000, and contain either two or more temporal periods. Furthermore, the types of change

scenarios detected within these datasets are also diverse (e.g., water area and densely built-up area). This diversity allows a more thorough evaluation of the robustness and effectiveness of the model.

2) *Standard Metrics*: Some quantitative evaluation indices derived from the confusion matrix were used in the study. To assess the quality of the generated DI maps, the receiver operating characteristic (ROC) curve and the precision–recall (PR) curve, along with their respective areas under the curves, namely, the area under ROC curve (AUR) and the area under PR curve (AUP) [50], were used. In addition, additional metrics, including overall accuracy (OA), kappa coefficient (KC), and $F1$ score ($F1$), were adopted to effectively evaluate the binary change maps.

3) *Parameter Settings*: There are several important parameters in our method. For the number of superpixels N , empirical evidence suggests that a typical 300×400 resolution image is adequately represented by 200 superpixels [30]. In accordance with previous work [17], N is set to 400 in this study. The initialization for the variables and parameters in the self-representation learning is set as $\{\mathbf{Z}_i = \mathbf{J}_i = \mathbf{E}_i = \mathbf{0}\}_{i=1}^k$, $\{\mathbf{Y}_i^1 = \mathbf{Y}_i^2 = \mathbf{0}\}_{i=1}^k$, $\mu^0 = 10^{-6}$, $\mu_{\max} = 10^{-6}$, and $\rho = 2.5$. For the two parameters σ_1 and σ_2 in background suppression stage, we empirically set $\sigma_1 = 0.15$ and $\sigma_2 = 0.4$. The optimal values of the crucial balancing parameters λ , γ , and β in (9) and η and θ in (20) are selected through experiments based on different datasets. A detailed analysis of their sensitivity is presented in Section IV-D.

B. Comparison With SOTAs

To validate the superiority of our proposed model, we conducted a thorough comparison with some traditional classical CD methods and some recently proposed advanced methods. These comparative approaches mainly include the following aspects.

- 1) *LR Operator* [10]: In this study, we treat it as the baseline method for performance comparison. The FCM algorithm is used to generate the final binary change map.
- 2) *Neighborhood-Ratio (NR) Operator* [51]: It utilizes local neighborhood information to weaken speckle noise, and we set the neighborhood size to 3×3 based on the suggested value in the original paper.
- 3) *Wavelet Fusion (WF)* [33]: It employs the wavelet transform technique for generating fused DI maps, effectively leveraging the complementary information between the mean-ratio image and the log-ratio image.
- 4) *NSCT-NLM* [34]: A fusion strategy based on the non-subsampled contourlet transform (NSCT) algorithm is proposed to fuse the results of the slow feature analysis with those of the LR operator, and the nonlocal means (NLMs) filtering algorithm is used to suppress the speckle noise.
- 5) *CWNN* [52]: It uses convolutional wavelet neural networks (CWNNs) to classify the changed and unchanged pixels. A virtual sample generation strategy is used to generate samples for CWNN training. The FCM algorithm is used to generate virtual training samples.
- 6) *MSDC-SAM* [53]: Multiscale self-attention (SAM) deep clustering is proposed for unsupervised difference representation learning. K -means++ clustering is used to generate a supervised signal to guide the learning of the model.
- 7) *RUSACD* [54]: A methodology for multiscale superpixel reconstruction is developed to produce the initial DI map, and a generative adversarial network is used to augment the pseudo-label training samples. In addition, CWNN is trained to categorize ambiguous samples.
- 8) *DDNet* [55]: It proposes the dual-domain network (DDNet), which combines spatial and frequency features from the discrete cosine transform, and designs a multi-region convolution module for noise reduction.
- 9) *LANTNet* [56]: It proposes a layer attention-based noise-tolerant network (LANTNet) that uses a layer attention module to adaptively weight convolutional features and a noise-tolerant loss function to mitigate the impact of noisy labels.
- 10) *CAMixer* [57]: It presents a transformer-like architecture that combines self-attention with shift convolution to capture global and local features, and uses a gating mechanism in the feed-forward network to enhance feature transformation and mitigate speckle noise.
- 11) *M₂HG* [58]: It is a graph-guided method using a heterogeneous graph with multilevel connections and a composite random walk matrix, applying a binary classification algorithm to generate the final binary change map.
- 12) *WBANet* [59]: It utilizes wavelet-based self-attention blocks with discrete wavelet transform (DWT) and inverse DWT to preserve information while enhancing local context.
- 13) *LRMD* [17]: It is a superpixel CD method that uses an LRMD strategy to generate the initial DI map, and an unsupervised linear classification refinement module is employed to refine the uncertain samples.

Among the competing algorithms, the first four belong to the traditional methods based on mathematical operators, while the rest are feature learning algorithms. They are all implemented using the source codes provided by the authors with default parameters. Considering that they are all designed for bitemporal images, we first compare their performance on six bitemporal datasets.

To evaluate the efficacy of S^3CD in change information extraction, we first conduct a detailed comparison of DI maps generated by competing algorithms. Notably, end-to-end deep learning methods inherently lack DI map outputs due to their integrated architecture. Therefore, we exclude such algorithms and focus exclusively on comparable approaches with explicit DI generation capabilities. Given the diversity of change types and environmental complexity, robust algorithms must: 1) accurately highlight genuine changes under heterogeneous perturbations and 2) effectively suppress unstable background noise. These capabilities serve as critical criteria for DI quality assessment.

Fig. 7 shows the visual comparison of the DI maps generated by the different methods. On some highly noisy datasets,

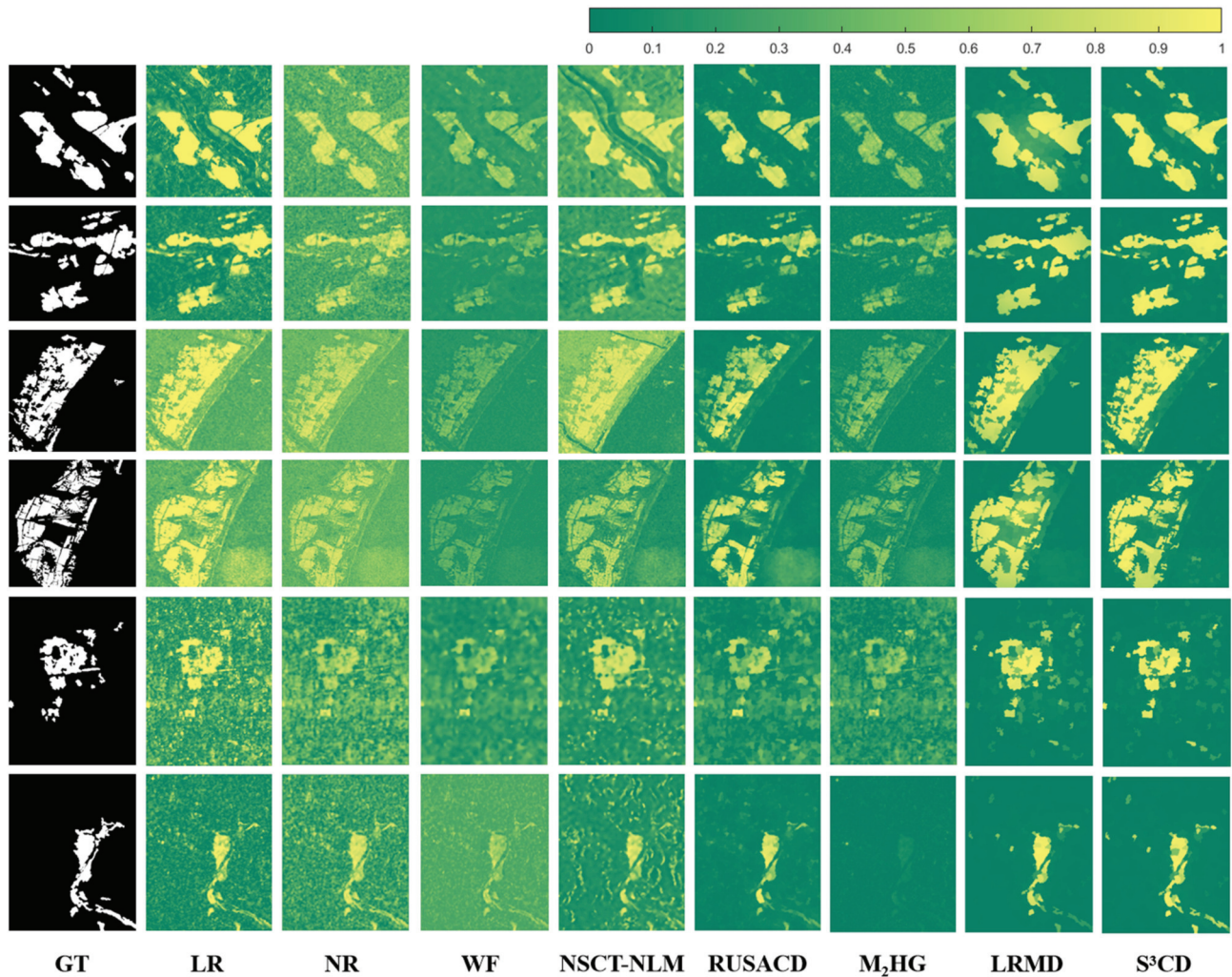


Fig. 7. Visualization of DI maps generated by different methods on six bitemporal datasets. From top to bottom, they correspond to datasets I–VI, respectively. From left to right, they correspond to ground truth and the generated DI maps by LR, NR, WF, NSCT-NLM, RUSACD, M_2 HG, LRMD, and proposed S^3 CD, respectively.

such as dataset IV (with severe sea clutter) and VI (affected by complex scatter interference from mountains), some methods (e.g., NR and WF) are difficult to effectively suppress coherent speckle noise, resulting in scattered background areas. This not only destroys the continuity of the changes, but also seriously affects their integrity. At the same time, the separability between the changed areas and the background in the DI maps is very low, and the boundaries are blurred. In contrast, RUSACD, LRMD, and the proposed S^3 CD show superior performance. For the building dataset V, the difficulty of extracting changed areas increases significantly due to the complex scattering interference from the surrounding dense buildings. However, the LRMD and the proposed S^3 CD algorithms perform well on such data, and are able to effectively separate the changed buildings from the complex background. Specifically, the LRMD algorithm has achieved certain results in terms of sparsity constraints on the changing structures, but is prone to the problem of local highlighting, which is particularly evident in datasets I and III. On the other hand, the S^3 CD algorithm considers both the sparsity and low-rank constraints,

resulting in good stationarity within the changing region and effectively avoiding the occurrence of small-scale highlighting regions. This feature not only improves the quality of DI maps, but also facilitates the subsequent binary segmentation processing.

Fig. 8 illustrates the ROC and PR curves of the DI maps generated by the different methods, accompanied by the corresponding AUR and AUP listed in Table II. The results show that the proposed S^3 CD exhibits the best performance on most of the datasets. Its average AUR and AUP reach 0.976 and 0.898, respectively, which are improved by 1.52% and 3.83%, respectively, compared to the next best method. This is attributed to the excellent structural feature learning capability of the low-rank self-representation model, as well as the fine-tuning of the results by the classification refine strategy. Specifically, the low-rank self-representation model is able to deeply mine the potential low-rank structure in the data, effectively capturing the key features of the changed regions. Meanwhile, by introducing the robust background suppression and label projection learning mechanism, it not

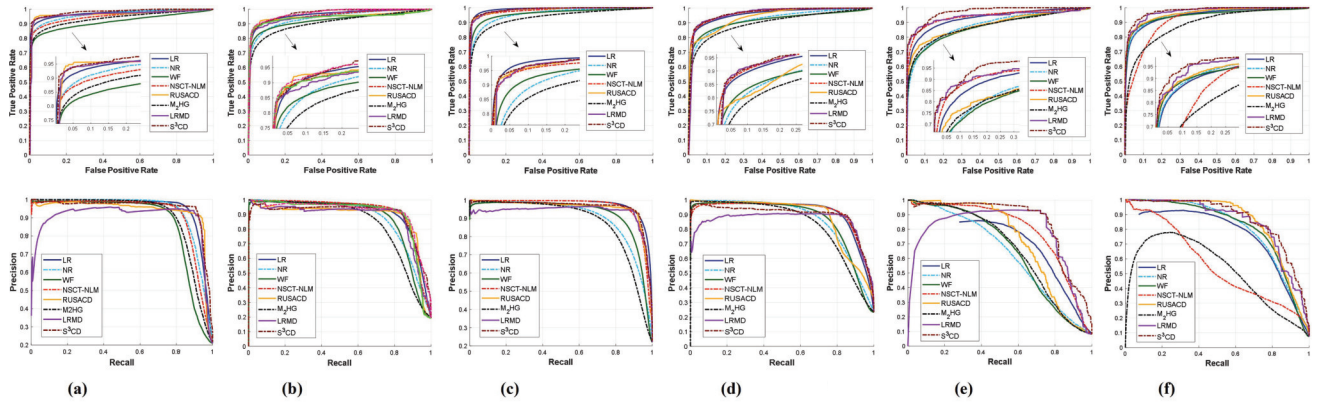


Fig. 8. ROC curves and PR curves of the DI maps generated by different methods. From left to right, they correspond to datasets I–VI, respectively. In each of the ROC plots, we have a zoomed-in view on one of the segments to show it more clearly. (a) Dataset I. (b) Dataset II. (c) Dataset III. (d) Dataset IV. (e) Dataset V. (f) Dataset VI.

TABLE II

COMPARISON WITH SOTA METHODS OF DI MAPS ON SIX BITEMPORAL DATASETS USING THE METRICS AUR AND AUP, WHERE $\uparrow$ DENOTES LARGER IS BETTER. THE BEST AND SECOND BEST RESULTS ARE SHOWN IN BOLD AND UNDERLINED, RESPECTIVELY

Datasets / Metrics	Dataset I		Dataset II		Dataset III		Dataset IV		Dataset V		Dataset VI		Average	
	AUR $\uparrow$	AUP $\uparrow$	AUR $\uparrow$	AUP $\uparrow$	AUR $\uparrow$	AUP $\uparrow$	AUR $\uparrow$	AUP $\uparrow$	AUR $\uparrow$	AUP $\uparrow$	AUR $\uparrow$	AUP $\uparrow$	AUR $\uparrow$	AUP $\uparrow$
LR [10]	0.9716	0.3406	<u>0.9612</u>	0.5718	0.9906	0.3286	0.9598	0.3501	0.9275	0.4564	0.9488	0.6922	0.9600	0.4566
NR [51]	0.9647	<u>0.9363</u>	0.9413	0.8775	0.9571	0.9050	0.9237	0.8647	0.8825	0.6319	0.9564	0.8069	0.9376	0.8370
WF [33]	0.9139	0.8768	0.9235	0.8723	0.9608	0.9218	0.9173	0.8646	0.8779	0.6513	0.9473	0.8116	0.9223	0.8331
NSCT-NLM [34]	0.9519	0.9176	0.9574	<u>0.8876</u>	0.9827	<u>0.9530</u>	<u>0.9661</u>	0.9153	0.9401	<u>0.7951</u>	0.9180	0.5591	0.9527	0.8378
RUSACD [54]	0.9759	0.9198	0.9448	0.8628	0.9757	0.9401	0.9414	0.8806	0.8861	0.7161	0.9623	<u>0.8403</u>	0.9491	<u>0.8600</u>
M ₂ HG [58]	0.9377	0.9034	0.9089	0.8380	0.9346	0.8787	0.9007	0.8333	0.8752	0.6648	0.8903	0.5143	0.9079	0.7721
LRMD [17]	<u>0.9639</u>	0.8940	0.9533	0.8696	0.9789	0.9291	0.9618	0.8518	<u>0.9391</u>	0.7469	<u>0.9676</u>	0.8312	0.9608	0.8568
S³CD	0.9817	0.9472	0.9671	0.9000	<u>0.9868</u>	0.9546	0.9676	<u>0.8912</u>	0.9705	0.8446	0.9789	0.8521	0.9760	0.8983

only accurately extracts the invariant background, but also effectively suppresses the saliency of the background noise and pseudo-changes. This strategy effectively improves the classification accuracy of boundary uncertain superpixels and further enhances the continuity of changes.

Fig. 9 presents binary change maps from all comparative methods, where white and black denote changed and unchanged regions, respectively, while red and cyan indicate omission and commission errors. Visually, the proposed S³CD demonstrates robust detection performance across both small- and large-scale datasets, exhibiting cleaner backgrounds with minimal clutter and more complete/coherent change boundaries. Although all methods perform satisfactorily on simpler datasets (I–III), with CWNN, LRMD, and S³CD showing particularly low error rates, most exhibit significant performance degradation on complex scenarios (datasets IV–VI). Particularly for urban environments (dataset V), conventional algorithms struggle to distinguish changes amid strong background interference, generating severe false alarms. By contrast, S³CD accurately extracts authentic structural changes even in cluttered settings, validating its superior robustness.

The comparative results of the quantitative evaluation are summarized in Table III, where the top two of each evaluation indicator are highlighted in bold and underlined, respectively. As shown in Table III, the proposed method outperforms the other comparative methods on the vast majority of evaluation metrics. In particular, on datasets V and VI, the KC values of

S³CD increased by 2.29% and 2.77%, respectively, compared to the suboptimal results. In addition, the average values of the method on the three evaluation indicators are improved by 0.27%, 1.80%, and 1.93%, respectively, compared to the suboptimal results. This indicates that the proposed method has higher robustness and reliability compared to the comparative methods.

In order to further verify the robustness and scalability of the proposed method, additional experiments were conducted on multitemporal datasets VII–IX. Fig. 10 visually displays the DI maps and classification results on these datasets, with specific quantitative evaluation results listed in Table IV. The results show that our method also exhibits good performance under multitemporal conditions. Even on the complex and challenging building dataset VIII, the proposed method is still able to produce robust detection results with a KC of 0.73. This proves that the proposed method can still maintain high adaptability and robustness in the face of more complex and diverse practical application scenarios, demonstrating great potential for monitoring continuous changes.

C. Ablation Study

The proposed S³CD includes two core steps: multiview self-representation learning and classification refinement. Among them, the first stage is subdivided into the joint low-rank and sparse constraints modeling and robust background suppression. To verify the effectiveness of each module for the

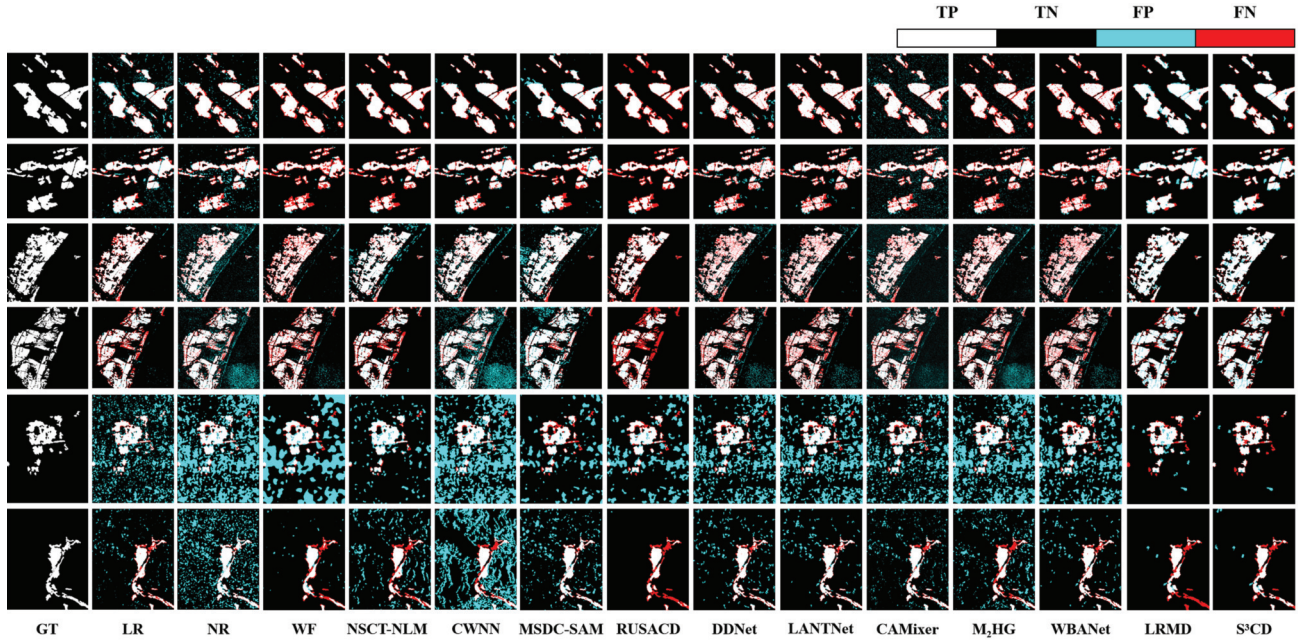


Fig. 9. Visualization of the binary change maps obtained using different methods on six bitemporal datasets. From top to bottom, they correspond to datasets I–VI, respectively. From left to right, they correspond to ground truth and the binary change maps generated by LR, NR, WF, NSCT-NLM, CWNN, MSDC-SAM, RUSACD, DDNet, LANTNet, CAMixer, M_2HG , WBANet, LRMD, and proposed S^3CD , respectively. In the maps, white represents the true positive (TP), black is the true negative (TN), cyan indicates the false positive (FP), and red stands as the false negative (FN).

TABLE III

COMPARISON WITH SOTA METHODS OF DIS ON SIX BITEMPORAL DATASETS USING THE METRICS OA, $F1$, AND KC, WHERE $\uparrow$ DENOTES LARGER IS BETTER. THE BEST AND SECOND BEST RESULTS ARE SHOWN IN BOLD AND UNDERLINED, RESPECTIVELY

Datasets / Metrics	Dataset I			Dataset II			Dataset III			Dataset IV			Dataset V			Dataset VI			Average		
	OA $\uparrow$	F1 $\uparrow$	KC $\uparrow$	OA $\uparrow$	F1 $\uparrow$	KC $\uparrow$	OA $\uparrow$	F1 $\uparrow$	KC $\uparrow$	OA $\uparrow$	F1 $\uparrow$	KC $\uparrow$	OA $\uparrow$	F1 $\uparrow$	KC $\uparrow$	OA $\uparrow$	F1 $\uparrow$	KC $\uparrow$	OA $\uparrow$	F1 $\uparrow$	KC $\uparrow$
Mathematical operator based methods																					
LR [10]	0.9417	0.8689	0.8316	0.9442	0.8537	0.8192	0.9602	0.9037	0.8788	0.9331	0.8414	0.7996	0.8102	0.4348	0.3549	0.9559	0.7181	0.6943	0.9242	0.7701	0.7297
NR [51]	0.9476	0.8727	0.8398	0.9239	0.8004	0.7533	0.9049	0.8027	0.7406	0.8683	0.7473	0.6595	0.6426	0.2901	0.1772	0.8324	0.4538	0.3837	0.8532	0.6611	0.5923
WF [33]	0.9499	0.8678	0.8374	0.9194	0.7444	0.6995	0.9364	0.8404	0.8012	0.9164	0.8025	0.7502	0.7220	0.3308	0.2294	<u>0.9637</u>	0.7233	0.7043	0.9013	0.7182	0.6703
NSCT-NLM [34]	0.9608	0.8989	0.8748	0.9329	0.7974	0.7586	0.9468	0.8830	0.8487	0.9348	0.8454	0.8047	0.9321	0.6833	0.6476	0.8978	0.4861	0.4331	0.9342	0.7657	0.7279
Feature learning based methods																					
CWNN [52]	0.9646	0.9122	0.8901	0.9442	0.8444	0.8107	0.9454	0.8779	0.8427	0.8402	0.7174	0.6107	0.6217	0.2788	0.1630	0.7143	0.2772	0.1768	0.8384	0.6513	0.5823
MSDC-SAM [53]	0.9601	0.9044	0.8792	0.9376	0.8229	0.7855	0.9407	0.8729	0.8344	0.9163	0.8332	0.7777	0.8857	0.5166	0.4586	0.9240	0.6347	0.5959	0.9274	0.7641	0.7218
RUSACD [54]	0.9486	0.8620	0.8311	0.9284	0.7772	0.7367	0.9509	0.8771	0.8468	0.9025	0.7429	0.6867	0.8449	0.4582	0.3865	0.9633	0.6799	0.6625	0.9231	0.7328	0.6917
DDNet [55]	0.9524	0.8811	0.8515	0.9252	0.7835	0.7391	0.9194	0.8067	0.7560	0.9067	0.7894	0.7297	0.7755	0.3779	0.2877	0.9414	0.6720	0.6407	0.9034	0.7184	0.6674
LANTNet [56]	0.9566	0.8885	0.8618	0.9273	0.7865	0.7437	0.9294	0.8306	0.7862	0.9078	0.7831	0.7252	0.7472	0.3483	0.2516	0.9580	<u>0.7489</u>	<u>0.7262</u>	0.9043	0.7309	0.6824
CAMixer [57]	0.9235	0.8226	0.7738	0.8978	0.7374	0.6740	0.8856	0.7570	0.6825	0.8716	0.7392	0.6543	0.8177	0.4188	0.3388	0.9473	0.6965	0.6680	0.8905	0.6952	0.6319
M_2HG [58]	0.9408	0.8468	0.8106	0.9205	0.7674	0.7204	0.9166	0.8078	0.7545	0.8782	0.7477	0.6675	0.6905	0.3120	0.2056	0.9057	0.5163	0.4670	0.8753	0.6663	0.6042
WBANet [59]	0.9520	0.8744	0.8451	0.9277	0.7823	0.7404	0.9204	0.8040	0.7546	0.9070	0.7847	0.7262	0.7782	0.3760	0.2861	0.9567	0.7323	0.7088	0.9070	0.7256	0.6768
LRMD [17]	<u>0.9694</u>	<u>0.9286</u>	<u>0.9091</u>	<u>0.9546</u>	<u>0.8820</u>	<u>0.8538</u>	<u>0.9694</u>	<u>0.9304</u>	<u>0.9108</u>	<u>0.9416</u>	0.8739	<u>0.8359</u>	<u>0.9672</u>	<u>0.7954</u>	<u>0.7776</u>	0.9623	0.6860	0.6675	<u>0.9607</u>	<u>0.8493</u>	<u>0.8257</u>
S^3CD	0.9717	0.9313	0.9135	0.9556	0.8828	0.8547	0.9672	0.9261	0.9050	0.9426	<u>0.8738</u>	0.8366	0.9731	0.8204	0.8060	0.9703	0.7693	0.7539	0.9634	0.8673	0.8450

TABLE IV

RESULTS OF QUANTITATIVE EVALUATION ON MULTITEMPORAL DATASETS USING THE PROPOSED S^3CD

Dataset / Metrics	OA $\uparrow$	F1 $\uparrow$	Kappa $\uparrow$
VII	0.9639	0.9089	0.8856
VIII	0.9361	0.7667	0.7312
IX	0.9727	0.9225	0.9060

algorithm framework, we performed ablation experiments on both bitemporal and multitemporal datasets.

Table V presents the quantitative evaluation results of the baseline model and its improved variants. The results demonstrate that the standalone LRR-based joint constraint model outperforms the robust background suppression approach

alone. Furthermore, after integrating the DI map fusion strategy, the classification performance of the model is further enhanced compared to single-processing methods, achieving a 2.02% improvement in the average KC over the standalone self-representation model. Moreover, with the inclusion of the classification refinement module, the model performance is further boosted, yielding an additional 1.26% increase in the average KC value. These results indicate that the simultaneous adoption of DI map fusion and classification refinement yields optimal model performance.

To further validate the impact of different constraint terms on model performance, Fig. 11 illustrates the T-SNE-based visualization of sample classification results. We can see that traditional RPCA-based methods (using either low-rank constraint $\|L\|_*$ or sparsity constraint $\|S\|_1$, even with Laplacian regularization) exhibit substantial interclass overlap and

TABLE V

ABLATION STUDY EXPERIMENT ON NINE DATASETS. BASE: MEANS USING THE STANDALONE JOINT LOW-RANK AND SPARSE CONSTRAINTS MODEL, BD: REPRESENTS THE ROBUST BACKGROUND DETECTION, AND RF: REPRESENTS THE CLASSIFICATION REFINEMENT MODULE

Metrics	Base.	BD.	RF.	I	II	III	IV	V	VI	VII	VIII	IX	Average
OA $\uparrow$	✓	×	×	0.9656	0.9466	0.9656	0.9361	0.9493	0.9685	0.9537	0.9344	0.9677	0.9542
	×	✓	×	0.9209	0.9170	0.9269	0.8898	0.9378	0.9668	0.9544	0.9157	0.9687	0.9331
	✓	✓	×	0.9698	0.9504	0.9656	0.9394	0.9658	0.9691	0.9590	0.9361	0.9701	0.9584
	✓	✓	✓	0.9717	0.9556	0.9672	0.9426	0.9731	0.9703	0.9629	0.9437	0.9727	0.9622
F1 $\uparrow$	✓	×	×	0.9203	0.8569	0.9192	0.8563	0.7108	0.7578	0.8782	0.7536	0.9061	0.8391
	×	✓	×	0.8291	0.7863	0.8385	0.7901	0.7072	0.7443	0.8812	0.6895	0.9132	0.7977
	✓	✓	×	0.9283	0.8655	0.9208	0.8658	0.8014	0.7657	0.8927	0.7667	0.9194	0.8584
	✓	✓	✓	0.9313	0.8828	0.9261	0.8738	0.8204	0.7693	0.9089	0.7951	0.9225	0.8701
KC $\uparrow$	✓	×	×	0.8983	0.8242	0.8974	0.8154	0.6831	0.7414	0.8500	0.7179	0.8867	0.8127
	×	✓	×	0.7782	0.7308	0.7913	0.7166	0.6744	0.7270	0.8533	0.6430	0.8940	0.7565
	✓	✓	×	0.9074	0.8353	0.8988	0.8267	0.7779	0.7495	0.8674	0.7312	0.9017	0.8329
	✓	✓	✓	0.9135	0.8547	0.9050	0.8366	0.8060	0.7539	0.8856	0.7483	0.9060	0.8455

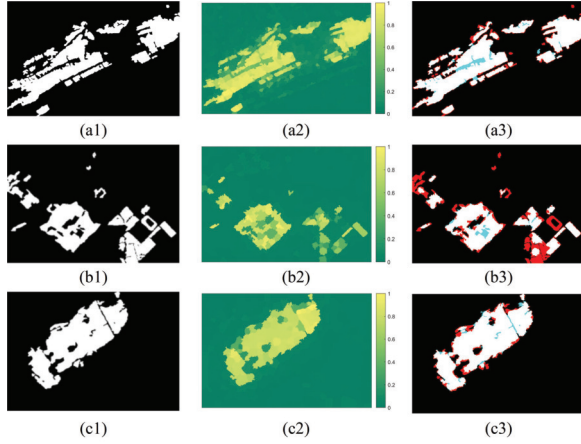


Fig. 10. Classification results on three multitemporal datasets using the proposed S^3CD . (a1)–(c1) Correspond to the ground truth of datasets VII–IX, respectively. (a2)–(c2) Correspond to the DI maps of datasets VII–IX, respectively. (a3)–(c3) Correspond to the binary change maps of datasets VII–IX, respectively.

blurred decision boundaries. The low-rank self-representation model integrated with Laplacian regularization significantly improves the separability between changed and unchanged regions, with most change points being effectively distinguished. The joint low-rank and sparsity constraints further refine the feature space, resulting in a well-defined linear decision plane and optimal separation between true changes and background artifacts. These findings conclusively verify that the multiconstraint joint optimization strategy enhances feature discriminability and achieves separation of changes more effectively.

D. Parameter Analysis

This section will analyze the sensitivity of the important balance parameters λ , γ , and β in (9) and η and θ in (20). In order to evaluate their impact on the classification accuracy, the relationship between KC and its values is analyzed.

The parameter λ governs the sparsity of $\mathbf{Z}$, constraining change regions to exhibit significant yet bounded differences.

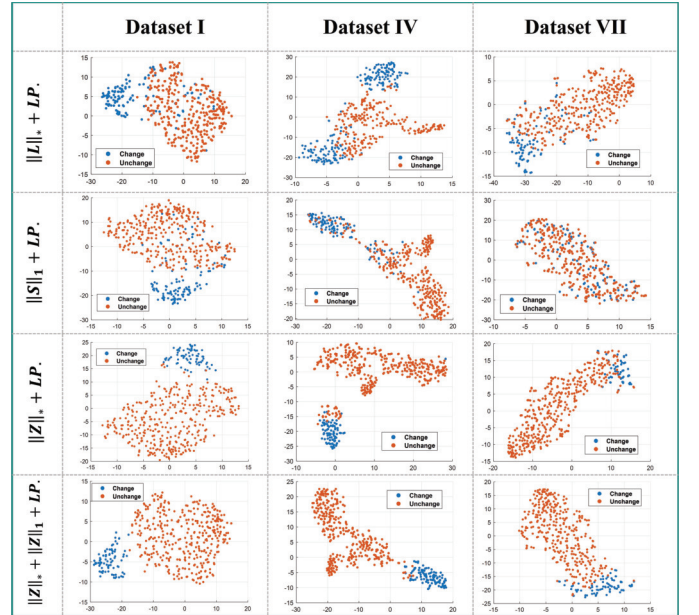


Fig. 11. T-SNE visualization of classification results with different constraint terms on datasets I, IV, and VII.

Fig. 12(a) shows the change of KC as λ increases from 0.2 to 2. It can be seen that for most of the datasets, λ takes the value of 0.4–0.8 with high KC values. And as λ continues to increase, there is a significant drop in model performance. The main reason is that as λ increases, sparsity has a greater impact on the optimization process, and the model will pay more attention to the sparse structure or outliers in the data, while relatively neglecting the overall control, leading to a decline in the performance of the model. To achieve stable performance, we set λ to 0.5 in the study.

The parameter γ mainly controls the sparsity of the error term. From Fig. 12(b), it can be seen that the KC of the different datasets have different trends with the variation of γ . For datasets I, II, and VI, the performance of the model remains stable within the range of value 0.6–2. For datasets III, IV, and V which are heavily affected by noise interference,

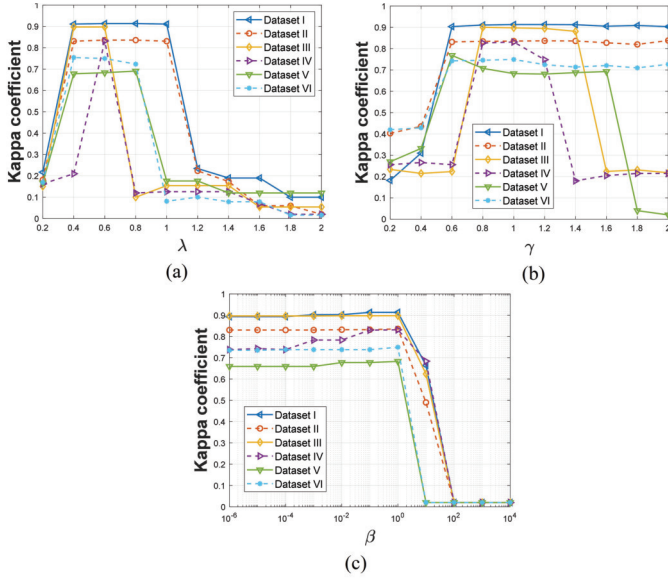


Fig. 12. Sensitivity analysis of the parameters λ , γ , and β . From left to right are the KC curves with different values of: (a) parameter λ , (b) parameter γ , and (c) parameter β .

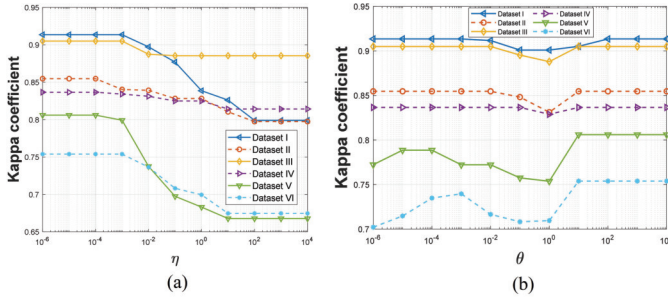


Fig. 13. Sensitivity analysis of the parameters (a) η and (b) θ .

the performance of the model decreases significantly between 1.2 and 1.6. To maintain high precision in all datasets, we set $\gamma = 1$ in the study.

The parameter β is used mainly to balance the hyper-Laplacian regularization term. As can be seen in Fig. 12(c), the model shows satisfactory performance and behaves smoothly when β varies over a wide range from 10^{-6} to 1. This means that in this range the model is not sensitive to the specific value of β and has strong robustness. When $\beta > 1$, the performance of the model drops significantly, indicating that too large a value of β destroys the regularization effect of the model. For simplicity, we set β to 1 in the study.

The parameter η primarily governs the intensity of the regularization term in nonlinear ridge regression. As shown in Fig. 13(a), an excessively large value of η will excessively suppress the complexity of the model, weaken the nonlinear expression capacity of the kernel space mapping. This excessive regularization impedes the model's ability to capture intricate variation patterns, consequently leading to diminished KC values. In order to suppress overfitting and retain sufficient nonlinear discriminative ability, this article sets η to 10^{-4} .

The parameter θ primarily governs the strength of the Laplacian regularization term. As demonstrated in Fig. 13(b),

TABLE VI
COMPARISON OF THE TIME COSTS FOR DIFFERENT PROCESSES OF THE PROPOSED FRAMEWORK ON DATASETS I, IV, AND VII

Datasets	N	t_1	t_2	t_3	t_{total}
I (286×298)	200	0.12	0.01	0.02	0.15
	400	0.20	0.02	0.02	0.24
	800	0.48	0.04	0.18	0.70
	1600	2.77	0.12	8.27	11.16
	3200	21.26	0.67	36.95	58.88
VII (635×422)	200	0.13	0.01	0.02	0.16
	400	0.17	0.01	0.02	0.20
	800	0.39	0.04	0.41	0.84
	1600	2.58	0.15	10.15	12.88
	3200	25.10	0.63	44.72	70.45
IV (1001×1001)	200	0.14	0.02	0.02	0.18
	400	0.21	0.02	0.02	0.25
	800	0.52	0.05	0.36	0.93
	1600	2.95	0.14	7.46	10.55
	3200	24.54	0.75	48.07	73.36

across the extensive parameter range of $\theta \in [10^{-6}, 10^4]$, the KC exhibits remarkable stability with fluctuation amplitudes below 5%, underscoring the model's high robustness to θ variations while confirming the spatial regularization's efficacy in noise suppression. Although a minor KC reduction emerges as θ approaches 1, this attenuation does not substantially disrupt significance distribution patterns. To circumvent potential performance degradation near $\theta = 1$, this study systematically employs $\theta = 10^2$ as the optimal configuration.

E. Complexity Analysis

The main computational process of the proposed S^3 CD consists of three parts: 1) optimization of the self-representation learning model; 2) execution of robust background detection and DI maps fusion; and 3) optimization of the classification refinement model. Given a change feature tensor $\mathcal{X} = [\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_k] \in \mathbb{R}^{d \times n \times k}$, it contains k feature views, each of which contains n feature vectors, and each vector has d -dimensional features. For each feature view, the matrices to be iteratively optimized include $\mathbf{Z}$, $\mathbf{J}$, $\mathbf{E}$, $\mathbf{Y}_1$, and $\mathbf{Y}_2$. Specifically, updating $\mathbf{Z}$ involves SVD and matrix subtraction, which has a computational complexity of $\mathcal{O}(dn^2)$ ($D \ll N$). To improve computational efficiency, the LADMAP algorithm provides an acceleration scheme for SVD that uses the Lanczos procedure to calculate the leading singular values and singular vectors, avoiding successive matrix-matrix multiplications. Therefore, its computational complexity is reduced to $\mathcal{O}(rn^2)$, where r represents the predicted rank of $\mathbf{Z}_i^{t+1}$ in (21). Updating $\mathbf{E}$ involves matrix multiplication and subtraction with a computational complexity of $\mathcal{O}(dn^2)$. Similar to updating $\mathbf{Z}$, it can also be reduced to $\mathcal{O}(rn^2)$ by utilizing Lanczos method. Updating $\mathbf{J}$ requires $\mathcal{O}(n^2)$, which only involves matrix addition. Therefore, the overall computational complexity of Algorithm 1 is $\mathcal{O}(krn^2)$. In the second phase, the computation of boundary connectivity necessitates pairwise comparisons between each superpixel and all others, resulting in a quadratic computational complexity of $\mathcal{O}(n^2)$. Regarding the third phase, given n_1 high-confidence samples participating in the learning pro-

cess, the closed-form solution (21) involves matrix inversion operations, thereby yielding a cubic computational complexity of $\mathcal{O}(n^3)$.

Table VI lists the time consumption of each of step of the proposed S^3 CD on datasets I, IV, and VII with different numbers of superpixels N . The experiment was executed under MATLAB 2020b on an Intel i5-10400 CPU with 32 GB of RAM. It is obvious that in all the datasets studied, the number of superpixels N has a significant effect on the computation time in the each step. Specifically, as N increases, the computation time of the first and third step shows a significant increasing trend, reflecting the direct contribution of N to the computational complexity, which is consistent with our previous complexity analysis results.

V. CONCLUSION

In this study, we propose a novel self-representation learning approach termed S^3 CD within a multiview framework, tailored for CD in SAR imagery. This method constructs the change feature tensor through multiple feature views, each encapsulating distinct structural information from the original observations to enrich the comprehensive data description. Simultaneously, the coefficient matrix undergoes both low-rank and sparse regularization, adeptly capturing local spatial structures while maintaining global consistency among homogeneous change objects. In addition, a robust background detection technique is employed to further suppress the speckle noise and pseudo-changes thereby significantly improving the quality of the DI. Furthermore, an unsupervised feature projection-based classification refinement module is incorporated to strengthen superpixel affiliation. Extensive evaluations across nine real-world datasets confirm the superior performance of our method compared to cutting-edge CD techniques, particularly showcasing its resilience in complex change scenarios.

Notably, our method exhibits impressive robustness not only on bitemporal datasets but also on multitemporal ones. Consequently, from an application perspective, the next step of research will focus on exploring the potential value of utilizing time series data to extend the work of S^3 CD. In addition, while our method leverages underlying structural features effectively, it currently excludes deep features. Thus, future investigations will actively pursue the integration of high-level features derived from deep networks with our existing methods, aiming to further elevate the performance of the algorithm.

VI. APPENDIX OPTIMIZATION FOR ALGORITHM 1

The optimization problem in (9) can be solved iteratively by the LADMAP algorithm. In the iterative procedure, the variables are updated alternately by minimizing the augmented Lagrangian function (11), with other variables fixed.

1) *Update $\mathbf{Z}_i$* : Based on (11), updating $\mathbf{Z}_i$ is equivalent to minimizing the following subproblem:

$$\begin{aligned} \mathbf{Z}_i^{t+1} &= \arg \min_{\mathbf{Z}_i} \mathcal{L}(\mathbf{Z}_i, \mathbf{J}_i^t, \mathbf{E}_i^t, \mathbf{Y}_1^t, \mathbf{Y}_2^t, \mu) \\ &= \arg \min_{\mathbf{Z}_i} \|\mathbf{Z}_i\|_* + \frac{\mu}{2} \left\| \mathbf{Z}_i - \mathbf{J}_i^t + \frac{1}{\mu} \mathbf{Y}_2^t \right\|_F^2 \end{aligned}$$

$$+ \frac{\mu}{2} \left\| \mathbf{X}_i - \mathbf{X}_i \mathbf{Z}_i^t - \mathbf{E}_i^t + \frac{1}{\mu} \mathbf{Y}_1^t \right\|_F^2 + \beta \text{tr}(\mathbf{Z}_i \mathbf{L}_i \mathbf{Z}_i^T). \quad (24)$$

Note that the above equation does not have a closed-form solution. For optimization purposes, an intermediate quadratic term is introduced

$$\begin{aligned} q(\mathbf{Z}_i, \mathbf{J}_i^t, \mathbf{E}_i^t, \mathbf{Y}_1^t, \mathbf{Y}_2^t, \mu) &= \frac{\mu}{2} \left\| \mathbf{Z}_i - \mathbf{J}_i^t + \frac{1}{\mu} \mathbf{Y}_2^t \right\|_F^2 \\ &\quad + \frac{\mu}{2} \left\| \mathbf{X}_i - \mathbf{X}_i \mathbf{Z}_i^t - \mathbf{E}_i^t + \frac{1}{\mu} \mathbf{Y}_1^t \right\|_F^2 \\ &\quad + \beta \text{tr}(\mathbf{Z}_i \mathbf{L}_i \mathbf{Z}_i^T). \end{aligned} \quad (25)$$

Then, the following closed-form solution for (19) can be derived

$$\begin{aligned} \mathbf{Z}_i^{t+1} &= \arg \min_{\mathbf{Z}_i} \|\mathbf{Z}_i\|_* + q(\mathbf{Z}_i, \mathbf{J}_i^t, \mathbf{E}_i^t, \mathbf{Y}_1^t, \mathbf{Y}_2^t, \mu) \\ &= \arg \min_{\mathbf{Z}_i} \|\mathbf{Z}_i\|_* + \langle \nabla_{\mathbf{Z}_i} q(\mathbf{Z}_i), \mathbf{Z}_i - \mathbf{Z}_i^t \rangle \\ &\quad + \frac{\mu\eta}{2} \|\mathbf{Z}_i - \mathbf{Z}_i^t\|_F^2 \\ &= \Gamma_{(\mu\eta)^{-1}}(\mathbf{Z}_i^t - \nabla_{\mathbf{Z}_i} q(\mathbf{Z}_i) / (\mu\eta)) \end{aligned} \quad (26)$$

where $\nabla_{\mathbf{Z}_i} q(\mathbf{Z}_i)$ denotes the partial differential of q with respect to $\mathbf{Z}_i$, and $\eta = \|\mathbf{X}_i\|_2^2$. $\Gamma_\epsilon(\mathbf{X}_i) = \mathbf{U} \Lambda_\epsilon[\Sigma] \mathbf{V}^T$, where $\mathbf{U} \Sigma \mathbf{V}^T$ is the singular value decomposition of matrix $\mathbf{X}_i$. $\Lambda_\epsilon(\mathbf{X}_i)$ is the contraction operator, which is defined as $\Lambda_\epsilon(\mathbf{X}_i) = \text{sgn}(\mathbf{X}_i) \cdot \max(|\mathbf{X}_i| - \epsilon, 0)$.

2) *Update $\mathbf{J}_i$* :

$$\begin{aligned} \mathbf{J}_i^{t+1} &= \arg \min_{\mathbf{J}_i \geq 0} \mathcal{L}(\mathbf{Z}_i^{t+1}, \mathbf{J}_i, \mathbf{E}_i^t, \mathbf{Y}_1^t, \mathbf{Y}_2^t, \mu) \\ &= \arg \min_{\mathbf{J}_i \geq 0} \lambda \|\mathbf{J}_i\|_1 + \frac{\mu}{2} \left\| \mathbf{J}_i - \mathbf{Z}_i^{t+1} - \frac{1}{\mu} \mathbf{Y}_2^t \right\|_F^2 \\ &= \max \left\{ \Gamma_{\frac{\lambda}{\mu}} \left(\mathbf{Z}_i^{t+1} + \frac{1}{\mu} \mathbf{Y}_2^t \right), 0 \right\}. \end{aligned} \quad (27)$$

3) *Update $\mathbf{E}_i$* :

$$\begin{aligned} \mathbf{E}_i^{t+1} &= \arg \min_{\mathbf{E}_i} \mathcal{L}(\mathbf{Z}_i^{t+1}, \mathbf{J}_i^{t+1}, \mathbf{E}_i, \mathbf{Y}_1^t, \mathbf{Y}_2^t, \mu) \\ &= \arg \min_{\mathbf{E}_i} \gamma \|\mathbf{E}_i\|_1 + \frac{\mu}{2} \left\| \mathbf{X}_i - \mathbf{X}_i \mathbf{Z}_i^{t+1} + \frac{1}{\mu} \mathbf{Y}_1^t - \mathbf{E}_i \right\|_F^2 \\ &= \Gamma_{\frac{\gamma}{\mu}} \left(\mathbf{X}_i - \mathbf{X}_i \mathbf{Z}_i^{t+1} + \frac{1}{\mu} \mathbf{Y}_1^t \right). \end{aligned} \quad (28)$$

4) *Update $\mathbf{Y}_1$ and $\mathbf{Y}_2$* :

$$\begin{aligned} \mathbf{Y}_1^{t+1} &= \arg \min_{\mathbf{Y}_1} \mathcal{L}(\mathbf{Z}_i^{t+1}, \mathbf{J}_i^{t+1}, \mathbf{E}_i^{t+1}, \mathbf{Y}_1, \mathbf{Y}_2^t, \mu) \\ &= \mathbf{Y}_1^t + \mu (\mathbf{X}_i - \mathbf{X}_i \mathbf{Z}_i^{t+1} - \mathbf{E}_i^{t+1}) \end{aligned} \quad (29)$$

$$\begin{aligned} \mathbf{Y}_2^{t+1} &= \arg \min_{\mathbf{Y}_2} \mathcal{L}(\mathbf{Z}_i^{t+1}, \mathbf{J}_i^{t+1}, \mathbf{E}_i^{t+1}, \mathbf{Y}_1^{t+1}, \mathbf{Y}_2, \mu) \\ &= \mathbf{Y}_2^t + \mu (\mathbf{Z}_i^{t+1} - \mathbf{J}_i^{t+1}). \end{aligned} \quad (30)$$

ACKNOWLEDGMENT

The authors gratefully acknowledge ESA for providing Sentinel-1 satellite data and Pasco Corporation for providing the TerraSAR-X satellite data. They would also like to thank the researchers for their generosity in sharing their source codes.

REFERENCES

- [1] D. Lu, "Change detection techniques," *Int. J. Remote Sens.*, vol. 25, no. 12, pp. 2365–2401, 2004.
- [2] J. Cihlar et al., "Change detection with synthetic aperture radar," *Int. J. Remote Sens.*, vol. 13, no. 3, pp. 401–414, 1992.
- [3] L. Giustarini, R. Hostache, P. Matgen, G. J.-P. Schumann, P. D. Bates, and D. C. Mason, "A change detection approach to flood mapping in urban areas using TerraSAR-X," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 4, pp. 2417–2430, Apr. 2013.
- [4] S.-W. Chen and M. Sato, "Tsunami damage investigation of built-up areas using multitemporal spaceborne full polarimetric SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 4, pp. 1985–1997, Apr. 2013.
- [5] P. Gamba, F. Dell'Acqua, and G. Lisini, "Change detection of multitemporal SAR data in urban areas combining feature-based and pixel-based techniques," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 10, pp. 2820–2827, Oct. 2006.
- [6] O. Yousif and Y. Ban, "Improving urban change detection from multitemporal SAR images using PCA-NLM," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 4, pp. 2032–2041, Apr. 2013.
- [7] R. Engeset et al., "Change detection and monitoring of glacier mass balance and facies using ERS SAR winter images over svalbard," *Int. J. Remote Sens.*, vol. 23, no. 10, pp. 2023–2050, 2002.
- [8] E. A. Lehmann, P. A. Caccetta, Z.-S. Zhou, S. J. McNeill, X. Wu, and A. L. Mitchell, "Joint processing of Landsat and ALOS-PALSAR data for forest mapping and monitoring," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 1, pp. 55–67, Jan. 2012.
- [9] R. A. Weismiller, S. J. Kristof, D. K. Scholz, P. E. Anuta, and S. A. Momin, "Change detection in coastal zone environments," *Photogramm. Eng. Remote Sens.*, vol. 43, no. 12, pp. 1533–1539, 1977.
- [10] Y. Bazi, L. Bruzzone, and F. Melgani, "Automatic identification of the number and values of decision thresholds in the log-ratio image for change detection in SAR images," *IEEE Geosci. Remote Sens. Lett.*, vol. 3, no. 3, pp. 349–353, Jul. 2006.
- [11] M. Hussain, D. Chen, A. Cheng, H. Wei, and D. Stanley, "Change detection from remotely sensed images: From pixel-based to object-based approaches," *ISPRS J. Photogramm. Remote Sens.*, vol. 80, pp. 91–106, Jun. 2013.
- [12] C. Marin, F. Bovolo, and L. Bruzzone, "Building change detection in multitemporal very high resolution SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 5, pp. 2664–2682, May 2015.
- [13] H. Zhang, M. Lin, G. Yang, and L. Zhang, "ESNet: An end-to-end superpixel-enhanced change detection network for very-high-resolution remote sensing images," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 1, pp. 28–42, Jan. 2023.
- [14] H. Chen, N. Yokoya, C. Wu, and B. Du, "Unsupervised multimodal change detection based on structural relationship graph representation learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, Dec. 2022, Art. no. 5635318.
- [15] 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.
- [16] B. Du, L. Ru, C. Wu, and L. Zhang, "Unsupervised deep slow feature analysis for change detection in multi-temporal remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 12, pp. 9976–9992, Dec. 2019.
- [17] W. Li, H. Wang, and P. Ma, "Spatial correlation-constrained low-rank modeling for SAR image change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 62, 2024, Art. no. 5201415.
- [18] G. Liu, Z. Lin, S. Yan, J. Sun, Y. Yu, and Y. Ma, "Robust recovery of subspace structures by low-rank representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 1, pp. 171–184, Jan. 2013.
- [19] Y. Zhang, Z. Jiang, and L. S. Davis, "Learning structured low-rank representations for image classification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2013, pp. 676–683.
- [20] M. Yin, J. Gao, and Z. Lin, "Laplacian regularized low-rank representation and its applications," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 3, pp. 504–517, Mar. 2016.
- [21] E. J. Candès et al., "Robust principal component analysis?," *J. ACM*, vol. 58, no. 3, pp. 1–37, 2011.
- [22] N. Vaswani, T. Bouwmans, S. Javed, and P. Narayanamurthy, "Robust subspace learning: Robust PCA, robust subspace tracking, and robust subspace recovery," *IEEE Signal Process. Mag.*, vol. 35, no. 4, pp. 32–55, Jul. 2018.
- [23] W. Lan et al., "Multiview subspace clustering via low-rank symmetric affinity graph," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 8, pp. 11382–11395, Aug. 2024.
- [24] Y. Tang, Y. Xie, and W. Zhang, "Affine subspace robust low-rank self-representation: From matrix to tensor," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 8, pp. 9357–9373, Aug. 2023.
- [25] J. Chen, S. Yang, H. Mao, and C. Fahy, "Multiview subspace clustering using low-rank representation," *IEEE Trans. Cybern.*, vol. 52, no. 11, pp. 12364–12378, Nov. 2022.
- [26] H. Gao, F. Nie, X. Li, and H. Huang, "Multi-view subspace clustering," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 4238–4246.
- [27] S. Luo et al., "Consistent and specific multi-view subspace clustering," in *Proc. AAAI Conf. Artif. Intell.*, vol. 32, no. 1, 2018, pp. 3730–3737.
- [28] X. Si, Q. Yin, X. Zhao, and L. Yao, "Consistent and diverse multi-view subspace clustering with structure constraint," *Pattern Recognit.*, vol. 121, Jan. 2022, Art. no. 108196.
- [29] L. Zhuang, H. Gao, Z. Lin, Y. Ma, X. Zhang, and N. Yu, "Non-negative low rank and sparse graph for semi-supervised learning," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 2328–2335.
- [30] W. Zhu, S. Liang, Y. Wei, and J. Sun, "Saliency optimization from robust background detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 2814–2821.
- [31] H.-C. Li, T. Celik, N. Longbotham, and W. J. Emery, "Gabor feature based unsupervised change detection of multitemporal SAR images based on two-level clustering," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 12, pp. 2458–2462, Dec. 2015.
- [32] S. Cui and M. Datcu, "Statistical wavelet subband modeling for multitemporal SAR change detection," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 4, pp. 1095–1109, Aug. 2012.
- [33] M. Gong, Z. Zhou, and J. Ma, "Change detection in synthetic aperture radar images based on image fusion and fuzzy clustering," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 2141–2151, Apr. 2012.
- [34] W. Li, X. Xiao, P. Xiao, H. Wang, and F. Xu, "Change detection in multitemporal SAR images based on slow feature analysis combined with improving image fusion strategy," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 15, pp. 3008–3023, 2022.
- [35] Y. Li, C. Peng, Y. Chen, L. Jiao, L. Zhou, and R. Shang, "A deep learning method for change detection in synthetic aperture radar images," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 8, pp. 5751–5763, Aug. 2019.
- [36] S. Saha, F. Bovolo, and L. Bruzzone, "Building change detection in VHR SAR images via unsupervised deep transcoding," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 3, pp. 1917–1929, Mar. 2021.
- [37] C. E. Woodcock, T. R. Loveland, M. Herold, and M. E. Bauer, "Transitioning from change detection to monitoring with remote sensing: A paradigm shift," *Remote Sens. Environ.*, vol. 238, Mar. 2020, Art. no. 111558.
- [38] G. Quin, B. Pinel-Puysségur, J.-M. Nicolas, and P. Loreaux, "MIMOSA: An automatic change detection method for SAR time series," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 9, pp. 5349–5363, Sep. 2014.
- [39] E. Colin Koeniguer and J.-M. Nicolas, "Change detection based on the coefficient of variation in SAR time-series of urban areas," *Remote Sens.*, vol. 12, no. 13, p. 2089, Jun. 2020.
- [40] T. T. Lê, J.-L. Froger, and D. Ho Tong Minh, "Multiscale framework for rapid change analysis from SAR image time series: Case study of flood monitoring in the central coast regions of Vietnam," *Remote Sens. Environ.*, vol. 269, Feb. 2022, Art. no. 112837.
- [41] W. Li, P. Ma, H. Wang, and C. Fang, "SAR-TSCC: A novel approach for long time series SAR image change detection and pattern analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 5203016.
- [42] M. Gong, T. Zhan, P. Zhang, and Q. Miao, "Superpixel-based difference representation learning for change detection in multispectral remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 5, pp. 2658–2673, May 2017.
- [43] N. Lv, C. Chen, T. Qiu, and A. K. Sangaiah, "Deep learning and superpixel feature extraction based on contractive autoencoder for change detection in SAR images," *IEEE Trans. Ind. Inform.*, vol. 14, no. 12, pp. 5530–5538, Dec. 2018.
- [44] Y. Lei, X. Liu, J. Shi, C. Lei, and J. Wang, "Multiscale superpixel segmentation with deep features for change detection," *IEEE Access*, vol. 7, pp. 36600–36616, 2019.
- [45] X. Zhang, Y. Zheng, J. Feng, and S. Gou, "SAR image change detection based on low rank matrix decomposition," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2012, pp. 6271–6274.
- [46] B. Hou, Y. Wang, and Q. Liu, "Change detection based on deep features and low rank," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 12, pp. 2418–2422, Dec. 2017.

- [47] W. Zhang, J. Li, F. Zhang, J. Sun, and K. Zhang, "Unsupervised change detection of multispectral images based on PCA and low-rank prior," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022.
- [48] Z. Lin, R. Liu, and Z. Su, "Linearized alternating direction method with adaptive penalty for low-rank representation," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 24, 2011, pp. 1–9.
- [49] A. Antelmi, G. Cordasco, M. Polato, V. Scarano, C. Spagnuolo, and D. Yang, "A survey on hypergraph representation learning," *ACM Comput. Surveys*, vol. 56, no. 1, pp. 1–38, Jan. 2024.
- [50] Y. Sun, L. Lei, L. Liu, and G. Kuang, "Structural regression fusion for unsupervised multimodal change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 4504018.
- [51] M. Gong, Y. Cao, and Q. Wu, "A neighborhood-based ratio approach for change detection in SAR images," *IEEE Geosci. Remote Sens. Lett.*, vol. 9, no. 2, pp. 307–311, Mar. 2012.
- [52] F. Gao, X. Wang, Y. Gao, J. Dong, and S. Wang, "Sea ice change detection in SAR images based on convolutional-wavelet neural networks," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 8, pp. 1240–1244, Aug. 2019.
- [53] H. Dong, W. Ma, L. Jiao, F. Liu, and L. Li, "A multiscale self-attention deep clustering for change detection in SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5207016.
- [54] X. Zhang, H. Su, C. Zhang, X. Gu, X. Tan, and P. M. Atkinson, "Robust unsupervised small area change detection from SAR imagery using deep learning," *ISPRS J. Photogramm. Remote Sens.*, vol. 173, pp. 79–94, Mar. 2021.
- [55] X. Qu, F. Gao, J. Dong, Q. Du, and H.-C. Li, "Change detection in synthetic aperture radar images using a dual-domain network," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022.
- [56] D. Meng, F. Gao, J. Dong, Q. Du, and H.-C. Li, "Synthetic aperture radar image change detection via layer attention-based noise-tolerant network," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022.
- [57] H. Zhang, Z. Lin, F. Gao, J. Dong, Q. Du, and H.-C. Li, "Convolution and attention mixer for synthetic aperture radar image change detection," *IEEE Geosci. Remote Sens. Lett.*, vol. 20, pp. 1–5, 2023.
- [58] J. Wang, F. Zeng, S. Niu, J. Zheng, and X. Jiang, "SAR image change detection via heterogeneous graph with multiorder and multilevel connections," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 17, pp. 11386–11401, 2024.
- [59] J. Xie, F. Gao, X. Zhou, and J. Dong, "Wavelet-based bi-dimensional aggregation network for SAR image change detection," *IEEE Geosci. Remote Sens. Lett.*, vol. 21, pp. 1–5, 2024.



Yinwei Li (Member, IEEE) received the bachelor's degrees in electronic information engineering from the University of Electronic Science and Technology of China, Chengdu, China, in 2009, and the Ph.D. degree in signal and information processing from the University of Chinese Academy of Sciences, Beijing, China, in 2014.

He was with Shanghai Radio Equipment Research Institute, Shanghai, China, from 2014 to 2019. Since 2019, he has been with the Terahertz Technology Innovation Research Institute, Terahertz Spectrum and Imaging Technology Cooperative Innovation Center, University of Shanghai for Science and Technology, Shanghai. He is currently with Shanghai Institute of Intelligent Science and Technology, Tongji University, Shanghai, and the School of Intelligent Emergency Management, University of Shanghai for Science and Technology, where he is also a Professor. His research interests include (inverse) synthetic aperture radar (SAR/ISAR) imaging, radar system design and signal processing, and terahertz radar imaging.



Yiming Zhu (Senior Member, IEEE) received the Ph.D. degree from The University of Tokyo, Tokyo, Japan, in 2008.

He is currently with Shanghai Key Laboratory of Modern Optical System and the Terahertz Technology Innovation Research Institute, University of Shanghai for Science and Technology, Shanghai, China, and the Terahertz Spectrum and Imaging Technology Cooperative Innovation Center, Shanghai. He is the Director of the Terahertz Spectrum and Imaging Technology Cooperative Innovation Center and the Terahertz Precision Biomedical Technology Overseas Expertise Introduction Center for Discipline Innovation, University of Shanghai for Science and Technology. Up to now, he has published more than 100 articles, including more than 40 articles in *Light: Science & Applications*, *Advanced Optical Materials*, *Applied Physics Letters*, *Optics Letters*, and *Optics Express* (top 5%), including five articles are selected as ESI articles. His research interests include terahertz technologies and applications, including terahertz devices, terahertz spectroscopy, imaging systems, and terahertz bio-applications.

Dr. Zhu is currently an Award Committee Member of the International Society of Infrared, Millimeter Wave and Terahertz (IRMMW-THz), a Council Member of the China Instrument and Control Society, a Council Member of the China Optical Engineering Society, and a Council Member of the Young Scientist Club of China Electronics Society. He is the Topic Editor of *Light: Advanced Manufacturing and PhotonIX*.



Weisong Li (Member, IEEE) received the M.S. degree in electronic and communication engineering from the University of Chinese Academy of Sciences, Beijing, China, in 2020, and the Ph.D. degree from the Key Laboratory of Information Science of Electromagnetic Waves (MoE), Fudan University, Shanghai, China, in 2024.

He is currently a Post-Doctoral Researcher with the Terahertz Technology Innovation Research Institute, University of Shanghai for Science and Technology, Shanghai. His research interests include

SAR image interpretation, multitemporal analysis, and change detection.



Haipeng Wang (Senior Member, IEEE) received the B.S. and M.S. degrees in mechanical and electronic engineering from Harbin Institute of Technology, Harbin, China, in 2001 and 2003, respectively, and the Ph.D. degree in environmental systems engineering from Kochi University of Technology, Kochi, Japan, in 2006.

He was a Visiting Researcher with the Graduate School of Information, Production and Systems, Waseda University, Fukuoka, Japan, in 2008. He is currently a Professor with the Key Laboratory of Electromagnetic Wave Information Science (MoE), Department of Communication Science and Engineering, School of Information Science and Engineering, Fudan University, Shanghai, China. His research interests include signal processing, SAR image processing and analysis, speckle statistics, and applications to forestry and oceanography, and machine learning and its applications to SAR images.

Dr. Wang has been a member of the Technical Program Committee of the IEEE Geoscience and Remote Sensing Symposium (IGARSS) since 2011. He was a recipient of the Dean Prize of the School of Information Science and Engineering, Fudan University, in 2009, 2017, and 2021 respectively. He serves as an Associate Editor for IEEE GEOSCIENCE AND REMOTE SENSING LETTERS.