ZHOU, C., HE, Z., DONG, J., LI, Y., REN, J. and PLAZA, A. 2025. Low-rank and sparse representation meet deep unfolding: a new interpretable network for hyperspectral change detection. *IEEE transactions on geoscience and remote sensing* [online], Early Access. Available from: https://doi.org/10.1109/TGRS.2025.3564996

Low-rank and sparse representation meet deep unfolding: a new interpretable network for hyperspectral change detection.

ZHOU, C., HE, Z., DONG, J., LI, Y., REN, J. and PLAZA, A.

2025

Low-Rank and Sparse Representation Meet Deep Unfolding: A New Interpretable Network for Hyperspectral Change Detection

Chengle Zhou, *Graduate Student Member, IEEE,* Zhi He, *Senior Member, IEEE,* Jian Dong, Yunfei Li, *Member, IEEE,* Jinchang Ren, *Senior Member, IEEE,* and Antonio Plaza, *Fellow, IEEE*

***Abstract*—Hyperspectral image change detection (HSI-CD) is a technique that intelligently checks the changed details in bi-temporal hyperspectral images (Bi-HSIs). Deep learning (DL), with the ability to model nonlinear changing features, has achieved promising results in HSI-CD, but the feature mining mechanism is unclear and the architecture design lacks transparency in such DL models. To alleviate this problem, this paper proposes a new low-rank and sparse representation-based deep unfolding network (LRSRNet) for HSI-CD. For feature mining mechanism, the LRSRNet adopts a low-rank and sparse sub-network (LRSnet) and a change detection sub-network (CDnet). The former is responsible for extracting low-rank features with valuable information and suppressing sparse features containing interference information, while the latter aims to obtain change information from low-rank features. For architecture design, the LRSnet formulates the HSI as a low-rank estimation, sparse estimation, and hyperspectral reconstruction in a low-rank and sparse model, and iteratively optimizes and updates the above sub-problems through deep networks. A new CDnet is designed as a concise convolutional architecture to extract change information from representative Bi-HSIs features. Experiments on three real datasets demonstrate the performance superiority of the proposed LRSRNet method over nine model-driven, data-driven, and model-data-joint-driven HSI-CD algorithms in both qualitative and quantitative evaluations. The proposed LRSRNet is available online: https://github.com/chengle-zhou/LRSRNet.**

***Index Terms*—Bi-temporal hyperspectral images, change detection, deep unfolding, low-rank and sparse representation.**

I. Introduction

HYperspectral image (HSI) is a fine-spectral Earth observation technique that provides diagnostic information on ground covers from visible light to shortwave infrared [1]–[3].

This work is supported by the National Natural Science Foundation of China under Grant No. 42271325, the National Key Research and Development Program of China under Grant No. 2020YFA0714103, Fundamental Research Funds for the Central Universities, Sun Yat-sen University under Grant No. 24lgqb002, and the Innovation Group Project of Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai) under Grant No. 311022018. (*Corresponding author: Zhi He.*)

Chengle Zhou, Zhi He, Jian Dong, and Yunfei Li are with the School of Geography and Planning, Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), Sun Yat-sen University, Guangzhou 510275 (e-mail: chengle_zhou@foxmail.com; hezh8@mail.sysu.edu.cn; dongj39@mail2.sysu.edu.cn; liyf18213483@163.com).

Jinchang Ren is with the School of Computer Sciences, Guangdong Polytechnic Normal University, Guangzhou, China. He is also with the National Subsea Centre, Robert Gordon University, AB10 7AQ Aberdeen, U.K. (e-mail: jinchang.ren@ieee.org).

Antonio Plaza is with the Hyperspectral Computing Laboratory, Department of Technology of Computers and Communications, Escuela Politecnica, University of Extremadura, E-10003 Caceres, Spain (e-mail: aplaza@unex.es).

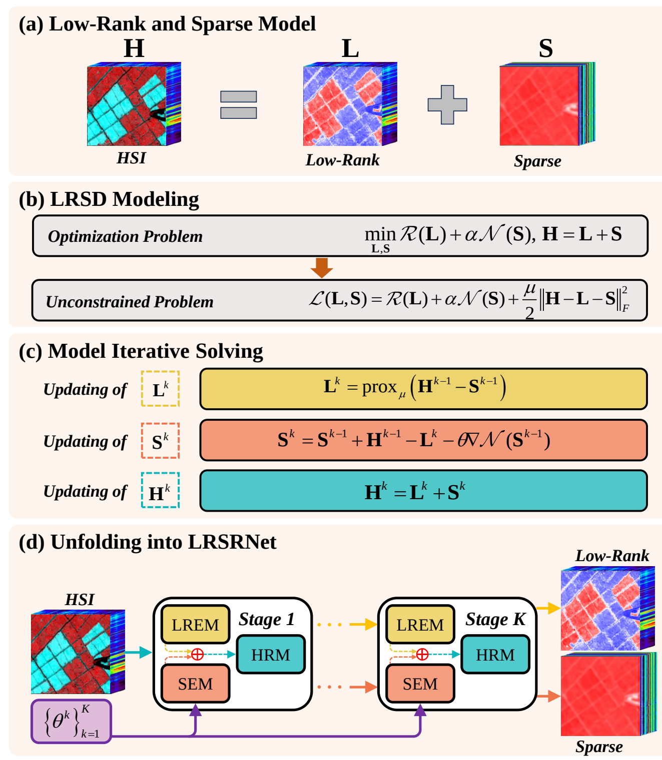


Fig. 1. Fundamental concept of the proposed LRSDNet. Note that the optimization procedure of the LRSD model is unfolded as a deep network.

Hyperspectral image change detection (HSI-CD) is a technical means to utilize HSIs of the same scene at different times [e.g., bi-temporal HSIs (Bi-HSIs)] to realize the perception of pixel, region, and scene change information, which has been widely used in precision agriculture [4], military reconnaissance [5] and disaster assessment [6].

The earliest HSI-CD methods can be roughly divided into three categories, namely, algebra-based methods, transformation-based methods, and classification-based methods [7]. Algebraic methods often use image difference information (e.g., spectral difference, absolute distance, and ratio) and image regression to obtain ground changes in Bi-HSIs scenes. For example, change vector analysis (CVA) determines the degree of coverage change based on the magnitude and direction of the change vector of Bi-HSIs [8], [9]. However, CVA cannot take into account the spectral variability caused by solar altitude, atmospheric conditions, and soil wetlands, resulting in serious pseudo-change pixels in the results (i.e., a high false alarm rate). The basic design of transformation-based methods is to project the original data to find a feature space that can linearly distinguish between changing and non-changing information [10]. Celik [11] employed principal component analysis (PCA) to obtain low-dimensional representative features of the original HSIs and used k-Means clustering to further obtain the change information in the scene. Nielsen [12] introduced an iterative reweighted MAD method based on canonical correlation analysis between Bi-HSIs. Wu et al. [13] proposed a slow feature analysis to capture the significance of changes between Bi-HSIs. Classification-based methods include post-classification and direct classification. The former respectively classifies the bi-temporal observation images and obtains change information through class difference analysis. The latter uses a classifier to perform binary classification on its difference features to obtain change results. Classic methods include support vector machine (SVM) [14], maximum likelihood algorithm [15] and k-nearest neighbor (k-NN) [16]. The above traditional HSI-CD methods can be summarized as empirical mathematical models (i.e., model-driven), which have the advantage of clearly defining the feature extraction mechanism and change detection rules in Bi-HSIs. That is, the mechanism is simple and clear. However, the above methods suffer from noise interference in practical applications, and there are problems such as insufficient feature utilization and difficulty in detecting subtle changes.

Recently, data-driven deep learning (DL) methods have achieved promising performance in the field of HSI-CD, thanks to the ability of deep networks to capture nonlinear changing characteristics [17]–[20]. Representative deep network architectures include recurrent neural networks (RNNs) [21], convolutional neural networks (CNNs) [22], and transformers [23]. Mou et al. [21] introduced a recurrent convolutional neural network for change detection in remote sensing images to learn a joint spectral-spatial-temporal feature representation framework. Chen et al. [24] proposed a deep Siamese convolutional multi-layer recurrent neural network to obtain change information in dual-temporal images. Shi et al. [25] designed a multi-path convolutional recurrent neural network to capture the multi-scale temporal-spatial-spectral variation characteristics in Bi-HSIs. The above RNNs architectures all use long short-term memory (LSTM) to associate the changing details of the dual-phase image scenes [26]. Du et al. [22] and Wang et al. [27] designed CNN-based HSI-CD architectures from the perspective of deep representation of slow features and endmember abundance, respectively. Ou et al. [28] proposed a CNN framework with compact band weighting and multi-scale spatial attention for HSI-CD. A multi-decision joint alignment framework using CNNs designed for HSI-CD in [29]. Since CNN, which only has the ability to summarize local features, has difficulty capturing global dependencies, the transformer architecture came into being. Li et al. [23] proposed a cross-band 2-D self-attention network to preserve the spectral characteristics in HSI-CD task. Wang et al. [30] developed a triple-branch transformer network based on parent and sibling attention for HSI-CD. Xiao et al. [31] proposed a novel selective transformer architecture that effectively addresses the bottleneck in feature extraction. Besides, convolutional transformer-based network architectures have also been widely investigated in HSI-CD applications [32], [33]. Although the above data-driven methods have achieved promising detection results in the HSI-CD task, there are three aspects that deserve further attention:

1) *Architecture Design*: Deep networks all use basic units such as convolutional layers and fully connected layers to build network architectures, ignoring experience-guided modeling.
2) *Feature Mining*: It relies on complex deep networks to abstract and represent the temporal-spatial-spectral features in Bi-HSIs gradually, ignoring the domain knowledge of image processing.
3) *Sample Desire*: Detection performance relies on iterative training of a large amount of labeled data, and the performance will be unsatisfactory when samples are limited.

More recently, the model-data-joint-driven approaches have been favored by scholars in the field of HSI processing due to the transparency of the model architecture and its lightweight characteristics [34]–[36]. For instance, Li et al. proposed a series of interpretable deep unfolding network schemes for anomaly detection problem [37], [38]. Nie et al. [39] proposed a deep unfolding network representation model for HSI classification task. Xiao et al. [40] introduced an efficient diffusion probabilistic model for super-resolution task, which shown remarkable reconstruction performance. Besides, Zhang et al. [41] introduced a spatial-spectral dual-image unfolding network for multispectral and hyperspectral fusion. In the HSI-CD task, some scholars have designed DL networks based on the dictionary learning mechanism for HSI-CD, which on the one hand enhances the transparency of the network structure and on the other hand improves the detection performance of the dictionary learning algorithm. Zhao et al. [42] designed a dictionary learning-guided deep interpretability network for HSI-CD. The fundamental concept is to assume that Bi-HSIs share an overcomplete dictionary and thus define scene attributes (i.e., changed or unchanged) based on the sparse difference. Qu et al. [43] introduced the multi-scale strategy into the dictionary learning network and proposed a multi-scale convolutional sparse coding-guided deep interpretable network for HSI-CD. The above literature provides successful cases and solutions of the model-data joint-driven paradigm in the field of HSI processing.

In this paper, we explore the unfolding network paradigm of coupling low-rank and sparse representation model with DL for HSI-CD. The motivation behind this work is that temporal-spatial-spectral feature extraction between Bi-HSIs is prone to over-detection and under-detection due to hyperspectral imaging being affected by many factors (e.g., atmospheric conditions and sensor jitter). Thus, it is necessary to introduce domain knowledge to guide feature extraction as shown in Fig. 1. In this work, we use the low-rank and sparse representation model to obtain low-frequency low-rank features and high-frequency sparse features from the raw HSI. It is emphasized that the low-rank features contain the intrinsic information of

HSI, while the sparse features belong to interference noise such as sunlight. Based on this fundamental consideration, a deep unfolding network based on low-rank and sparse representation (LRSRNet) for HSI-CD is proposed, which includes several key technical steps. Firstly, low-rank and sparse representation is utilized to decouple HSI into low-rank and sparse features. Then, an individually iterative scheme is introduced to update the low-rank features and sparse features. Next, a low-rank and sparse subnetwork (LRSnet) is proposed to learn low-rank estimation, sparse estimation, and hyperspectral reconstruction. Meanwhile, a K-step iterative scheme is applied to LRSnet to update the learnable low-rank and sparse features. Finally, a change detection subnetwork (CDnet) constructed by concise convolution is designed to extract change information from the last updated low-rank features.

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

1) To the best of our knowledge, low-rank and sparse representation are investigated for the first time using a deep unfolding network for the HSI-CD task, which are a new model-data joint-driven network structure for temporal-spatial-spectral feature capture in Bi-HSIs.
2) Low-rank and sparse feature learning are defined as two independent sub-problems. By expanding the iterative optimization update steps into a deep network framework, the time-consuming and complex matrix calculations are replaced by theory-guided neural networks.
3) A residual attention learning mechanism (RALM) with coordinate and spectral information considerations is designed in the sub-problem solving to fully perceive the semantic information of HSI.

Importantly, experiments on several real Bi-HSIs shown the performance superiority of the LRSRNet over nine model-driven, data-driven, and model-data-joint-driven HSI-CD approaches in both qualitative and quantitative evaluations.

The remainder of this paper is organized as follows. Section II provides a review of LRSR in HSI processing including change detection and other fields. Section III introduces the proposed LRSRNet method in detail. Section IV presents experimental analysis and discussions on Bi-HSIs datasets. Section V concludes the paper with some remarks and hints at plausible future research.

II. LRSR in HSI Processing

Due to the high correlation between spectral channels and spatial pixels, HSI have intrinsically low-rank and sparse (LRSR) structures [44]. The LRSR representation methods have confirmed to be powerful tools for HSI processing and are widely used in various HSI fields including denoising [45], classification [46], anomaly detection [47], and change detection [48]. For low-rank matrix recovery (LRMR) theory [49], assume that an observed image $\mathbf{Y}$ can be regarded as the combination of an ideal low-rank image $\mathbf{X}$ and a sparse noise matrix $\mathbf{E}$, which can be formulated as follows:

$$\mathbf{Y} = \mathbf{X} + \mathbf{E} \tag{1}$$

The low-rank matrix $\mathbf{X}$ and sparse noise matrix $\mathbf{S}$ can be recovered according to robust principal component analysis (RPCA) by solving the following optimization problem [49]:

$$\min \|\mathbf{X}\|_* + \lambda\|\mathbf{E}\|_1, \quad \mathbf{Y} = \mathbf{X} + \mathbf{E} \tag{2}$$

where $\|\cdot\|$ refers to the nuclear norm of a matrix, and λ represents a regularization parameter.

Zhou et al. [50] improved the RPCA model (2) by considering both sparse noise $\mathbf{E}$ and Gaussian random noise $\mathbf{N}$. Their observed model is

$$\mathbf{Y} = \mathbf{X} + \mathbf{E} + \mathbf{N} \tag{3}$$

and the corresponding optimization problem is as follows:

$$\min \|\mathbf{X}\|_* + \lambda\|\mathbf{E}\|_1, \quad \text{s.t.,} \quad \|\mathbf{Y} - \mathbf{X} - \mathbf{E}\|_\text{F}^2 \leq \delta \tag{4}$$

where δ is a constant related to the standard deviation of random noise $\mathbf{N}$. According to low-rank matrix decomposition, some change detection algorithms have been successfully developed. For instance, Zhao et al. [48] proposed a spectrally-spatially regularized low-rank and sparse decomposition for HSI-CD. Zhang et al. [51] introduced an unsupervised HSI-CD algorithm based on PCA and low-rank prior. Hou et al. [52] developed a low-rank deep feature-based CNN for HSI-CD. It is worth mentioning that more information about LRSR theory and applications can be found in [44].

III. Methodology

This section presents the fundamental concepts, design ideas, and solutions of the proposed LRSRNet method. The framework of the proposed method is shown in Figs 2-3. The proposed LRSRNet method will be introduced in six parts, including problem definition, model iterative solution scheme, deep unfolding network scheme, change detection network, and loss function.

A. Problem Definition

Given a HSI $\mathbf{H}$, it can be separated into low-rank features $\mathbf{L}$ and sparse features $\mathbf{S}$ by the physical model, which are defined as follows:

$$\mathbf{H} = \mathbf{L} + \mathbf{S}, \tag{5}$$

where $\mathbf{H}$, $\mathbf{L}$, and $\mathbf{S} \in \mathbb{R}^{M\times N\times C}$. M, N, and C are height, width and number of channels of the HSI. Note that $\mathbf{L}$ represents the intrinsic characteristics of HSI and $\mathbf{S}$ includes the noise caused by hyperspectral imaging.

According to the theory of low-rank and sparse decomposition [50], [53], the sparse features $\mathbf{S}$ are usually constrained to recover the low-rank features $\mathbf{L}$ that can be paired with the given $\mathbf{H}$, thus transforming the problem into:

$$\min_{\mathbf{L},\mathbf{S}} \text{rank}(\mathbf{L}) + \alpha\|\mathbf{S}\|_0 \quad s.t. \quad \mathbf{H} = \mathbf{L} + \mathbf{S}, \tag{6}$$

where α stands for a trade-off parameter and $\|\cdot\|_0$ is the l_0-norm recorded as the number of non-zero elements.

Due to low-rank functions and L_0 norms are non-convex and discontinuous, the solution of Eq. (6) is NP-hard. Therefore, many model designs employed nuclear norms and L_1

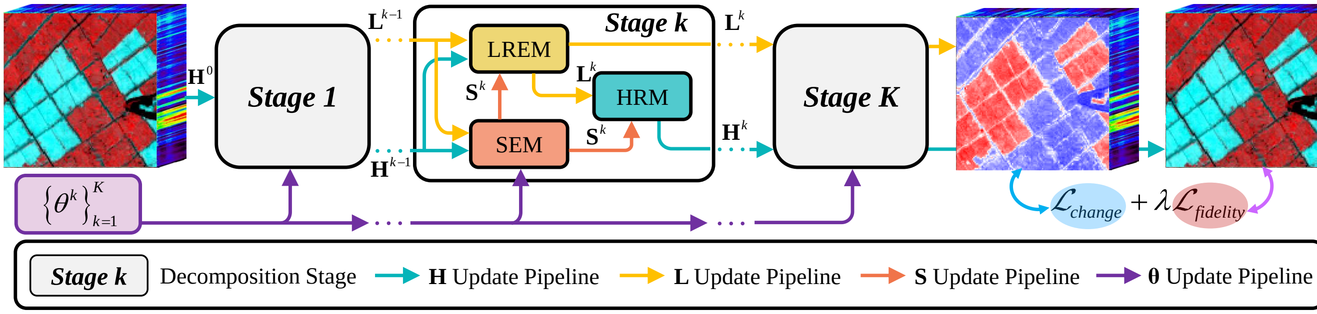


Fig. 2. Overall structure of the proposed LRSRNet for the HSI-CD task.

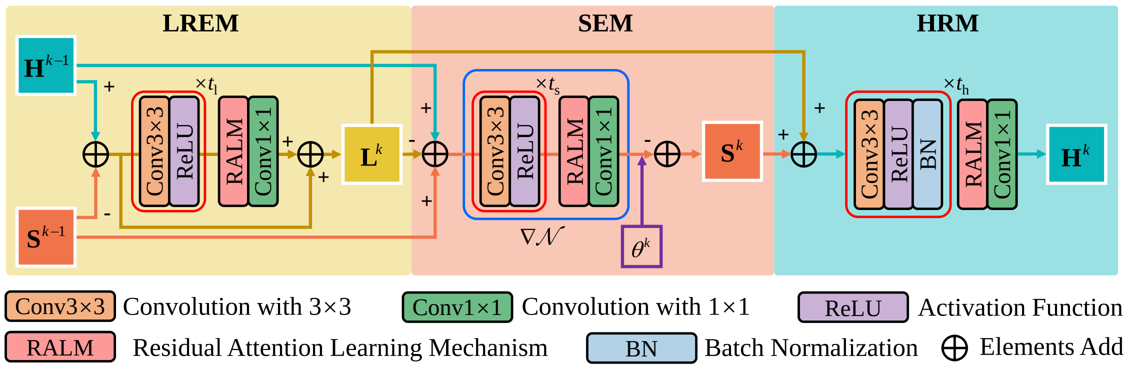


Fig. 3. Detailed network structure of the single stage from Fig. 2 in LRSRNet including Low-rank estimation module (LREM), Sparse estimation module (SEM), and hyperspectral reconstruction module (HRM).

norms to replace them according to the principal component pursuit. As such, Eq. (6) can be reformulated as:

$$\min_{\mathbf{L},\mathbf{S}} \|\mathbf{L}\|_* + \alpha\|\mathbf{S}\|_1 \quad s.t. \quad \mathbf{H} = \mathbf{L} + \mathbf{S}. \tag{7}$$

In complex hyperspectral scenes, low-frequency low-rank and high-frequency sparseness may differ in complexity and sparsity [54], and an alone norm or rank function is difficult to capture the actual constraints. Thus, $\mathcal{R}(\mathbf{L})$ and $\mathcal{N}(\mathbf{S})$ are introduced to constrain the prior knowledge of low-rank and sparse features, respectively. The expression is as follows:

$$\min_{\mathbf{L},\mathbf{S}} \mathcal{R}(\mathbf{L}) + \alpha\mathcal{N}(\mathbf{S}), \quad \mathbf{H} = \mathbf{L} + \mathbf{S}. \tag{8}$$

Furthermore, in order to avoid the complexity of augmented Lagrange multipliers in the process of updating variables, a more concise and intuitive L_2 norm is utilized to convert the constrained problem to an unconstrained problem [55], the expression is as follows:

$$\mathcal{L}(\mathbf{L},\mathbf{S}) = \mathcal{R}(\mathbf{L}) + \alpha\mathcal{N}(\mathbf{S}) + \frac{\mu}{2}\|\mathbf{H} - \mathbf{L} - \mathbf{S}\|_F^2, \tag{9}$$

where μ represents a penalty coefficient and $\|\cdot\|_F$ refers to the Frobenius norm (F-Norm). We emphasized that the low-rank and sparse features can be individually optimized in an iterative scheme based on Eq. (9).

B. Model Iterative Solution Scheme

Updating $\mathbf{L}^*$: To update the low-rank components, the sub-problem of solving $\mathbf{L}^*$ is formulated as:

$$\mathbf{L}^* = \arg\min_{\mathbf{L}} \mathcal{R}(\mathbf{L}) + \frac{\mu}{2}\|\mathbf{L} + \mathbf{S} - \mathbf{H}\|_F^2. \tag{10}$$

According to Eq. (7), traditional optimization-based methods often convert $\mathcal{R}(\mathbf{L})$ into $\|\mathbf{L}\|_*$, and then Eq. (10) can be degenerated into the sum of the nuclear norm and the L_2 norm. The analytical solution to this problem can be expressed as:

$$\mathbf{L}^* = \mathcal{O}_\varepsilon(\mathbf{H} - \mathbf{S}), \tag{11}$$

where $\mathcal{O}_\varepsilon(\cdot)$ refers to the singular value thresholding algorithm (SVT) [56] with threshold of ε. The solution of Eq. (11) involves singular value decomposition (SVD) and its functions on singular values. If a neural network is used to replace this process, it is necessary to perform SVD on each neural tensor in the forward propagation, which cannot avoid time-consuming and accuracy issues. To alleviate this problem, an initialization method according to the best rank approximation of SVD is developed in [57] and a technique with various rank constraints on sub-matrices is investigated in [58]. However, the former still suffers from accuracy issues, while the latter relies on manual tuning and ignores the inherent properties of the image.

Unlike the above works, Eq. (11) is degenerated into a constraint function $\mathcal{R}(\mathbf{L})$ instead of employing the nuclear norm and solving it using complex SVD. Importantly, a proximal operator $\mathrm{Prox}_\varepsilon(\cdot)$ is used to approximate the closed-form solution of the low-rank component, which can be defined as follows:

$$\mathbf{L}^* = \mathrm{Prox}_\varepsilon(\mathbf{H} - \mathbf{S}), \tag{12}$$

where the convolutional layer is used to construct $\mathrm{Prox}_\varepsilon(\cdot)$ to solve the optimization problem. We emphasize that the convolutional network can effectively eliminate the time consumption of complex matrix operations. Meanwhile, thanks to the nonlinear ability of the network, it can capture the deep semantic features of the image in a data-driven way.

Updating $\mathbf{S}^*$: Similarity, to update the sparse components, the sub-problem of solving $\mathbf{S}^*$ is expressed as:

$$\mathbf{S}^* = \arg\min_{\mathbf{S}} \alpha\mathcal{N}(\mathbf{S}) + \frac{\mu}{2}\left\|\mathbf{S}+\mathbf{L}-\mathbf{H}\right\|_F^2. \tag{13}$$

As shown in Eq. (7), traditional optimization methods employ the L_1 norm to constrain sparse images, but there are challenges in mapping soft thresholds into neural networks. Besides, the sparse constraints should be shown adaptability to the scene changes. Based on this, the goal of our work is to derive a simple and intuitive representation of the closed-form solution of Eq. (13). To solve this problem, the Taylor expansion is applied to $\mathcal{N}(\mathbf{S})$, and then the last $\mathcal{N}(\mathbf{S})$ at $\mathbf{S}^{k-1}$ can be approximated as:

$$\hat{\mathcal{N}}(\mathbf{S},\mathbf{S}^{k-1}) \leftarrow \frac{L_s}{2}\left\|\mathbf{S}-\mathbf{S}^{k-1}+\frac{1}{L_s}\nabla\mathcal{N}(\mathbf{S}^{k-1})\right\|_2^2 + C_n \tag{14}$$

$$C_n = -\frac{1}{2L_s}\left\|\nabla\mathcal{N}(\mathbf{S}^{k-1})\right\|_2^2 + \mathcal{N}(\mathbf{S}^{k-1}), \tag{15}$$

where L_s and C_n represent the Lipschitz constant of $\mathcal{N}(\mathbf{S})$ and constant, respectively. And substituting Eq. (14) into Eq. (13), the update formulate of $\mathbf{S}$ can be redefined as:

$$\begin{aligned}\mathbf{S}^* &= \arg\min_{\mathbf{S}} \alpha\mathcal{N}(\mathbf{S}) + \frac{\mu}{2}\left\|\mathbf{S}+\mathbf{L}-\mathbf{H}\right\|_F^2 \\ &= \arg\min_{\mathbf{S}} \frac{L_s}{2}\left\|\mathbf{S}-\mathbf{S}^{k-1}+\frac{1}{L_s}\nabla\mathcal{N}(\mathbf{S}^{k-1})\right\|_2^2 \\ &\qquad + \frac{\mu}{2}\left\|\mathbf{S}+\mathbf{L}-\mathbf{H}\right\|_F^2,\end{aligned} \tag{16}$$

which involves only the sum of two L_2 norms instead of the traditional optimization problem with L_1 norm constraints, and thus the solution process does not require help from traditional optimization algorithms or simulated soft-thresholding methods [59]. Succinctly, by taking the derivative of Eq. (16) and making its derivative equal to zero, the closed solution of the k-th step update $\mathbf{S}$ can be obtained as follows:

$$\begin{aligned}\mathbf{S}^k = \frac{\alpha L_s}{\alpha L_s+\mu}\mathbf{S}^{k-1} &+ \frac{\mu}{\alpha L_s+\mu}\left(\mathbf{H}^{k-1}-\mathbf{L}^k\right) \\ &- \frac{\alpha}{\alpha L_s+\mu}\nabla\mathcal{N}(\mathbf{S}^{k-1}),\end{aligned} \tag{17}$$

where all the three coefficients are constants. By assigning a learnable vector to each vector, Eq. (17) for updating the sparse matrix can be rewritten as:

$$\begin{aligned}\mathbf{S}^k =& \beta\mathbf{S}^{k-1} + (1-\beta)\left(\mathbf{H}^{k-1}-\mathbf{L}^k\right) - \theta\nabla\mathcal{N}(\mathbf{S}^{k-1}) \\ \beta =& \frac{\alpha L_s}{\alpha L_s+\mu}, \theta = \frac{\alpha}{\alpha L_s+\mu}.\end{aligned} \tag{18}$$

In our work, an end-to-end deep network satisfying the Lipschitz continuity assumption is employed to learn $\nabla\mathcal{N}(\cdot)$ functions instead of complex matrix operations (e.g., soft-thresholding algorithms). Thus, the reconstructed HSI for the k-th update can be obtained as follows:

$$\mathbf{H}^k = \mathbf{L}^k + \mathbf{S}^k. \tag{19}$$

C. Deep Unfolding Network Scheme

To reduce complex matrix operations, a low-rank and sparse representation subnetwork (LRSnet) is introduced to replace the low-rank and sparse decomposition process. As shown in Fig. 2, the input to the network is a single-temporal HSI $\mathbf{I} \in \mathbb{R}^{M\times N\times C}$, where M, N, and C are the height, width, and number of channels of the HSI, respectively, as well as the parameters to be updated are initialized to $\mathbf{H}^0 = \mathbf{I}$ and $\mathbf{S}^0 = 0$. These parameters are then subjected to K decomposition learning stages, each of which corresponds to an iterative low-rank and sparse matrix decomposition process simulating the model-driven approach of multiple iterations of the update operation. Specifically, the updated parameters $\mathbf{H}^{k-1}$ and $\mathbf{S}^{k-1}$ are passed to the k-th stage ($k = 1, 2, 3, ..., K$). The low-rank component $\mathbf{L}_k$, the sparse component $\mathbf{S}_k$, and the reconstructed $\mathbf{H}_k$ of the current stage are estimated through the low-rank estimation module (LREM), the sparse estimation module (SEM), and the hyperspectral reconstruction module (HRM).

1) Low-rank Estimation Module (LREM): As shown in Fig. 3, a LREM module is proposed to achieve stage learning and estimation of low-rank components in HSI. Inspired by previous deep unfolding network-based works [60], [61], a flexible residual structure $\mathrm{ProxNet}_\varepsilon(\cdot)$ is developed to model the approximation operator $\mathrm{Prox}_\varepsilon(\cdot)$ in Eq. (12), which can be formulated as:

$$\begin{aligned}\mathbf{L}^k &= \mathrm{ProxNet}_\varepsilon\left(\mathbf{H}^{k-1}-\mathbf{S}^{k-1}\right) \\ &= \mathbf{H}^{k-1}-\mathbf{S}^{k-1}+\mathcal{F}^k\left(\mathbf{H}^{k-1}-\mathbf{S}^{k-1}\right),\end{aligned} \tag{20}$$

where $\mathcal{F}^k(\cdot)$ is denoted as the convolutional group. The stacking of t_l layers is first performed by 3×3 Conv and ReLU, followed by connecting to the proposed RALM module (see Fig. 4) and 1×1 Conv, where ReLU is the activation function. And the RALM is a joint deep feature capture module including convolutional group constructed by 3×3 Conv, ReLU, and BN with L_r layer, followed by the spectral attention mechanism (SAM) and coordinate attention mechanism (CAM). The details of the SAM and CAM can be found in [62] and [63], respectively.

2) Sparse Estimation Module (SEM): Also as shown in Fig. 3, the updated low-rank $\mathbf{L}^k$, sparse $\mathbf{S}^{k-1}$, and hyperspectral reconstruction $\mathbf{H}^{k-1}$ are fed as inputs to the SEM. Importantly, β is set to 0.5 to process the three inputs equally, and it is further expressed as follows:

$$\mathbf{S}^k = \mathbf{S}^{k-1} + \mathbf{H}^{k-1} - \mathbf{L}^k - \theta\nabla\mathcal{N}(\mathbf{S}^{k-1}), \tag{21}$$

where θ is a learnable scalar for the parameter learning task. Note that a is independent, that is, θ^k in each reconstruction stage does not share parameters. Due to the fact that the CNN constructed with convolutional layers and ReLU activation functions conforms to the Lipschitz continuum lemma, a CNN architecture with the same structure as the LREM is used to simulate the $\nabla\mathcal{N}(\cdot)$ function. Furthermore, the difference feature between $\mathbf{H}^{k-1}$ and the updated $\mathbf{L}^k$ is utilized to strengthen the sparse features, and the updated representation of $\mathbf{S}^k$ can be defined as:

$$\mathbf{S}^k = \mathbf{S}^{k-1} + \mathbf{H}^{k-1} - \mathbf{L}^k - \theta^k\mathcal{Z}^k(\mathbf{S}^{k-1}), \tag{22}$$

where $\mathcal{Z}^k(\cdot)$ is a CNN with a similar structure to LREM, but its initial module consisting of 3×3 Conv and ReLu is stacked t_s layers.

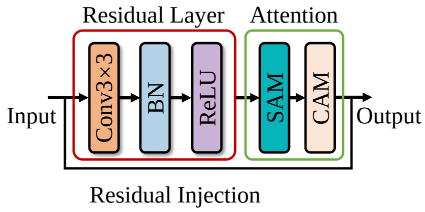


Fig. 4. Overview of the newly designed RALM in the proposed LRSRNet.

3) Hyperspectral Reconstruction Module (HRM): As shown in Fig. 3, we design the hyperspectral reconstruction module (HRM) to map the low-rank and sparse features into reconstructed HSI, which is composed of a neural network $\mathcal{J}^k(\cdot)$ that plays the role of going from decomposition task to reconstruction task. The expression is as follows:

$$\mathbf{H}^k = \mathcal{J}^k(\mathbf{L}^k + \mathbf{S}^k), \tag{23}$$

where $\mathcal{J}^k(\cdot)$ has a similar network structure to $\mathcal{F}^k(\cdot)$ and $\mathcal{Z}^k(\cdot)$. Considering the simplicity and lightness of the architecture, residual learning is removed and introduced into the batch norm (BN) layer. As shown in Fig. 3, the previous convolutional modules are stacked to construct the l_h layer.

D. Change Detection Network

Along with extracting the low-rank and sparse features of the HSIs, we apply the low-rank features containing the intrinsic attributes of the HSIs to the change detection task. As shown in Fig. 5, a concise Siamese change detection subnetwork (CDnet) is proposed for obtaining change information in Bi-HSIs. Firstly, we adopt the 2D convolution function with 3×3 kernel size, ReLU, BN, and (max pooling) MaxPool as the base unit to gradually capture the latent representation f_{out} of the change information, which can be formulated as follows:

$$\begin{cases} f_{\text{mid}} &= \text{BN}(\text{ReLU}(\text{Conv}_{3\times3}(\mathbf{L}^K(T_i)))) \\ f_{\text{out}} &= \text{Maxpool}(\text{BN}(\text{ReLU}(\text{Conv}_{3\times3}(f_{\text{mid}})))), \end{cases} \tag{24}$$

where $\mathbf{L}^K(T_i)$ stands for low-rank feature of HSI at T_i time (i = 1, 2). Once the low-rank latent representation of HSI is obtained, the matrix outer product is employed to define the difference features $\mathbf{\Gamma}(d)$ of Bi-HSI, as shown below:

$$\mathbf{\Gamma}(d) = \sum_{p\in P} f_{out}(T_1)^T f_{out}(T_2), \tag{25}$$

where $\mathbf{\Gamma}(d)$ refers to difference features of $\mathbf{L}^K(T_1)$ and $\mathbf{L}^K(T_1)$ at position d, and $f_{out}(T_1)$ and $f_{out}(T_2)$ are low-rank latent feature at T_1 and T_2 at a location $p \in P$. It is worth emphasizing that compared to performing difference operations on individual pixels [64] to obtain change information, the outer product can better combine neighboring information from patches at different times, rather than subtraction, ratio or log ratio of corresponding individual pixels.

Next, $\mathbf{\Gamma}(d)$ is reshaped to construct the vectorized feature representation $\mathbf{x}$. Moreover, $\mathbf{x}$ is further processed according to signed square root operation and L_2 normalization to improve the computational performance, which can be denoted as follows:

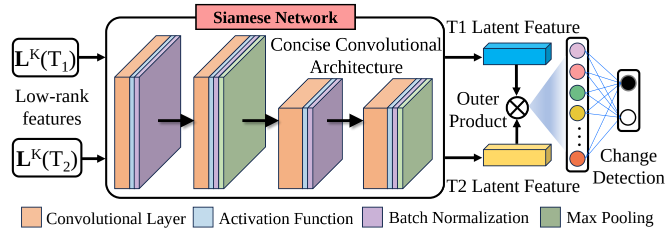


Fig. 5. Outline of Siamese network with simple convolutional architecture for change detection.

$$\mathbf{y} = \text{sign}(\mathbf{x})\sqrt{|\mathbf{x}|} \Big/ \left\|\text{sign}(\mathbf{x})\sqrt{|\mathbf{x}|}\right\|_2. \tag{26}$$

where sign$(\cdot)$ stands for a signed square root operation.

Finally, the standard feature vector $\mathbf{y}$ is fed into the fully connected layer and a softmax function is employed to obtain the output attribute probabilities (i.e., change and non-change), which can be described as follows:

$$\mathbf{s} \leftrightarrow \text{softmax}(\mathbf{y}) = e^{\mathbf{y}_i} \Big/ \sum_{c=1}^{C} e^{\mathbf{y}_i}. \tag{27}$$

E. Loss Function

In order to facilitate the effective joint learning of LRSnet and CDnet in the HSI-CD task, we design the reconstruction fidelity loss and binary change loss for LRSnet and CDnet, respectively. Specifically, the $\mathcal{L}_1$ loss function is employed to constrain the loss between $\hat{\mathbf{H}}$ (i.e., $\mathbf{H}^K$) and $\mathbf{H}$. and the cross-entropy loss function ($\mathcal{L}_{CE}$) is used to measure the difference between the predicted value of the change and the groundtruth, which can be defined as shown below:

$$\begin{aligned} \mathcal{L}_{\text{Total}} &= \mathcal{L}_{\text{change}} + \lambda\mathcal{L}_{\text{fidelity}} \\ &= \mathcal{L}_{\text{CE}} + \lambda\mathcal{L}_1, \end{aligned} \tag{28}$$

where λ is a trade-off parameter. Clearly, $\mathcal{L}_{\text{change}}$ is formulated as:

$$\mathcal{L}_{\text{CE}} = -\left[\mathbf{y}_{\text{true}} \ln(s(\mathbf{y})) + (1 - \mathbf{y}_{\text{true}}) \ln(s(1 - \mathbf{y}))\right], \tag{29}$$

where $s(\cdot)$ is softmax function shown in Eq. (23) and $\mathbf{y}_{\text{true}}$ is changed groundtruth in Bi-HSIs. $\mathcal{L}_1$ is expressed as follows:

$$\mathcal{L}_1 = \left|\hat{\mathbf{H}} - \mathbf{H}\right|. \tag{30}$$

Note that the specific impact of the value of λ will be discussed in detail later. For the implementation mechanism of LRSRNet in the program, a pseudo code is provided in Algorithm 1.

IV. Experimental Results and Analysis

A. Datasets and Evaluation Metrics

1) Datasets: The first performance validation scenario is the River dataset, which was collected using the Earth Observing-1 (EO-1) Hyperion sensor in a river area in Jiangsu Province, China, on May 3, 2013, and December 31, 2013, respectively, and consists of two HSIs with a total of 242 spectral bands

Algorithm 1: LRSRNet for HSI-CD

Input: $\mathbf{H} \in \mathbb{R}^{M\times N\times C}$ from Bi-HSIs
Output: Changed information $\mathbf{y}$ in Bi-HSIs

LRSnet for Bi-HSIs
Traditional Low-Rank and Sparse Model
$\mathbf{H} = \mathbf{L} + \mathbf{S}$
$\mathbf{L}^* = \arg\min_{\mathbf{L}} \mathcal{R}(\mathbf{L}) + \frac{\mu}{2}\|\mathbf{L}+\mathbf{S}-\mathbf{H}\|_F^2$
$\mathbf{S}^* = \arg\min_{\mathbf{S}} \alpha\mathcal{N}(\mathbf{S}) + \frac{\mu}{2}\|\mathbf{S}+\mathbf{L}-\mathbf{H}\|_F^2$
Its Deep Unfolding Network Scheme
initialization for $\mathbf{H}^0$, $\mathbf{S}^0 = 0$, etc.
# *Optimization With Deep Network in Training*
while $j < epoch$ **do**
  for *Stage* $k = 1, 2, ..., K$ **do**
    # Low-Rank Estimation Module (LREM)
    $\mathbf{L}^k = \mathbf{H}^{k-1} - \mathbf{S}^{k-1} + \mathcal{F}^k\left(\mathbf{H}^{k-1} - \mathbf{S}^{k-1}\right)$
    # Sparse Estimation Module (SEM)
    $\mathbf{S}^k = \mathbf{S}^{k-1} + \mathbf{H}^{k-1} - \mathbf{L}^k - \theta^k \mathcal{Z}^k(\mathbf{S}^{k-1})$
    # Hyperspectral Reconstruction Module (HRM)
    $\mathbf{H}^k = \mathcal{J}^k(\mathbf{L}^k + \mathbf{S}^k)$
  end
  # ***CDnet for Bi-HSIs***
  $f_{\text{mid}} = \text{BN}(\text{ReLU}(\text{Conv}_{3\times3}(\mathbf{L}^K(T_i))))$
  $f_{\text{out}} = \text{Maxpool}(\text{BN}(\text{ReLU}(\text{Conv}_{3\times3}(f_{\text{mid}}))))$
  $\mathbf{\Gamma}(d) = \sum_{p\in P} f_{out}(T_1)^T f_{out}(T_2)$
  # Calculate $\mathcal{L}_{\text{change}}$ and $\mathcal{L}_{\text{fidelity}}$ Joint Losses
  $\mathcal{L}_{\text{Total}} = \mathcal{L}_{\text{CE}} + \lambda\mathcal{L}_1$
end
Testing Stage for the LRSRNet
Import Bi-HSIs into LRSRNet, and obtain the change information $\mathbf{y}$ through the LRSnet and CDnet.

in each image. The single image has a spectral range of 0.4–2.5 μm, spatial and spectral resolutions of 30 m and 10 nm, respectively, and an image size of 463×241 pixels, with the primary change in coverage type in the image being riverine. False-color maps for the T_1 and T_2 HSIs of the river dataset, as well as changed groundtruth (GT), are provided in Fig. 6(a)-(c).

The second validation scenario is the Farmland dataset, which was also acquired by the EO-1 Hyperion sensor on May 3, 2006 and April 23, 2007 at a wetland agricultural area in Yancheng City, Jiangsu Province, China. The spectral range and spectral resolution of the single image are similar to those of the River dataset. The size of this dataset is 420×140 pixels and 154 spectral bands after removing the noise and water absorption bands are used for the experiments of performance testing in this work. The main coverage type that changes in HSIs is regular farmland. False color maps of the Farmland dataset along with GT are given in Fig. 6(d)-(f).

The third performance validation scenario is the Santa Barbara dataset, which was acquired by the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) sensor in the urban area of Santa Barbara, California, USA in 2013 and 2014, respectively. The spectral range, spectral resolution, and spatial resolution of the individual images are 0.4–2.5 μm, 10 nm, and 20 m. The dataset images are 984×740 pixels in size, contained 224 spectral bands, and the predominantly changing cover type is green vegetation. Fig. 6(g)-(k) presents false color images for T_1 and T_2 HSIs and GT between Bi-HSIs.

TABLE I
CONFUSION MATRIX FOR BINARY CLASSIFICATION. NOTE THAT TP, FN, FP, AND TN ARE TRUE POSITIVE, FALSE NEGATIVE, FALSE POSITIVE, AND TN: TRUE NEGATIVE, RESPECTIVELY.

Confusion Matrix		Prediction	
		Positive (Changed)	Negative (Unchanged)
GT	Positive	TP	FN
	Negative	FP	TN

2) Evaluation Metrics: To quantify the model performance in change detection, several generalized performance metrics based on binary confusion matrices, i.e., overall accuracy (OA), kappa coefficient (Kappa), precision (Precision), recall (Recall), and F1-score (F1). These performance metrics are formulated in detail as follows:

$$\text{OA} = \frac{\text{TP}+\text{TN}}{\text{TP}+\text{TN}+\text{FN}+\text{FP}}, \tag{31}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP}+\text{FP}}, \tag{32}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP}+\text{FN}}, \tag{33}$$

$$\text{F1} = \frac{2\times(\text{Precision}\times\text{Recall})}{\text{Precision}+\text{Recall}}, \tag{34}$$

$$\text{Kappa} = \frac{\text{OA}-P_e}{1-P_e}, \tag{35}$$

$$p_e = \frac{(\text{TP}+\text{FP})(\text{TP}+\text{FN})+(\text{TN}+\text{FN})(\text{TN}+\text{FP})}{(\text{TP}+\text{FN}+\text{FP}+\text{TN})^2}, \tag{36}$$

where TP denotes the number of pixels correctly identified as changing regions, TN denotes the number of pixels correctly identified as unchanging regions, FP denotes the number of pixels incorrectly identified as changing regions, and FN denotes the number of pixels incorrectly identified as unchanging regions (see Table I). The TP, TN, FP, and FN pixels are marked with white, black, red, and green color in the visualization results. The values of TP and TN have larger values indicating better detection performance of the model, while the opposite is true for FP and FN.

B. Competitor Methods and Implementation Details

In order to objectively evaluate the advancement of the proposed LRSRNet method in the HSI-CD task, we compared and analyzed the LRSRNet method with the well-known and state-of-the-art CD methods (including four model-driven, four data-driven, and one model-data joint-driven methods) on the above-mentioned Bi-HSIs data from 3 different scenarios.

- *Model-Driven Methods*: the CVA [9], k-means with PCA (PCAKM) [11], IR-MAD [12], and SVM [14] are added into our experimental comparison scheme. CVA is a

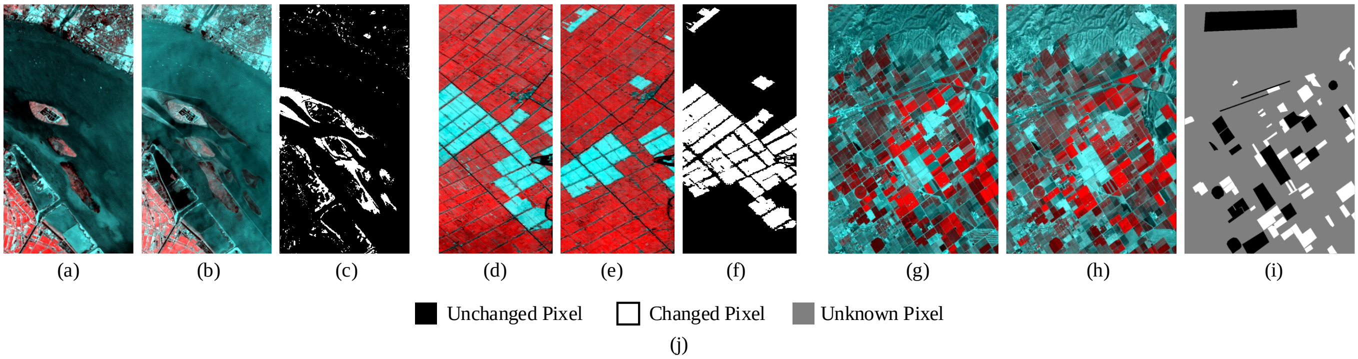


Fig. 6. False color images and GT for all datasets. (a)-(c) are false color images of T_1 and T_2 and GT on the River dataset. (d)-(f) are false color images of T_1 and T_2 and GT on the Farmland dataset. (g)-(i) are false color images of T_1 and T_2 and GT on the Santa Barbara dataset. (j) Pixel legend.

commonly used CD method that provides the intensity and direction of changes in Bi-HSIs. PCAKM is a binary classification result of difference features of Bi-HSIs using k-means. The basic idea of IR-MAD is to weight the iterations based on the cardinal distances on the basis of MAD. SVM is a radial basis function that is used as a kernel function and utilizes a five-fold cross-validation to determine its parameters.

- *Data-Driven Methods*: the cross-temporal interaction symmetric attention network (CSANet) [65], multiscale diff-changed feature fusion network (MSDFFN) [66], recurrent CNN (ReCNN) [67], and diff-feature contrast enhancement network (DCENet) [68] are adopt as the competitive CD methods in this work. CSANet aims at extracting and fusing the joint spatial-spectral-temporal features of HSI and enhancing the discrimination of changes. MSDFFN is concerned with learning the components of variation in the refinement between Bi-HSIs at different scales to improve feature representation. ReCNN is a combination of RNN and CNN used to learn joint spatial-spectral-temporal features of change features. DCENet combines spectral-spatial feature fusion with Siamese network learning to capture change information in Bi-HSIs.
- *Joint-Driven Method*: a multi-scale convolutional sparse coding (MSCSCNet)-guided deep interpretable network [43] is added into our comparative experiments. MSCSCNet adopts a model-driven network architecture design to capture shared and private sparse coefficients of different scales to determine the variation details between Bi-HSIs.

Moreover, we emphasize that the results of all competing methods that are added to the comparative analysis scheme are reproduced from results in the published literature. In the interest of fairness and objectivity, results are reported as the average of ten replicated experiments using the training and testing configuration in Table II.

In the proposed LRSRNet method, we randomly selected 3% labeled samples from River and Farmland datasets, and 0.2% labeled samples from Santa Barbara dataset as training set. The rest are all used as the test set. Note that the specific quantities can be found in Table II. We utilized the Adam optimizer with $\beta_1 = 0.9$, $\beta_2 = 0.999$ and $\varepsilon = 10^{-4}$ for parameter optimization. An exponential decay strategy is used to update the initial learning rate 2.0×10^{-4} for equal epoch steps, and the number of epochs is set to 200. The batch size used for training on all datasets is 128. The proposed LRSRNet method is trained by the Pytorch framework running in a Windows 11 environment with 32 GB RAM and an NVidia RTX 3060 GPU. Note that a detailed hyperparameter analysis will be given later in the discussion section.

TABLE II
NUMBER OF TRAINING AND TEST SAMPLES FOR VARIOUS DATASETS.

Dataset	Training Set			Testing Set		
	UCH	CH	Total	UCH	CH	Total
River	2400	1200	3600	99485	8498	107983
Farmland	1342	549	1891	43381	17728	61109
Santa Barbara	860	251	1111	79558	51883	131441

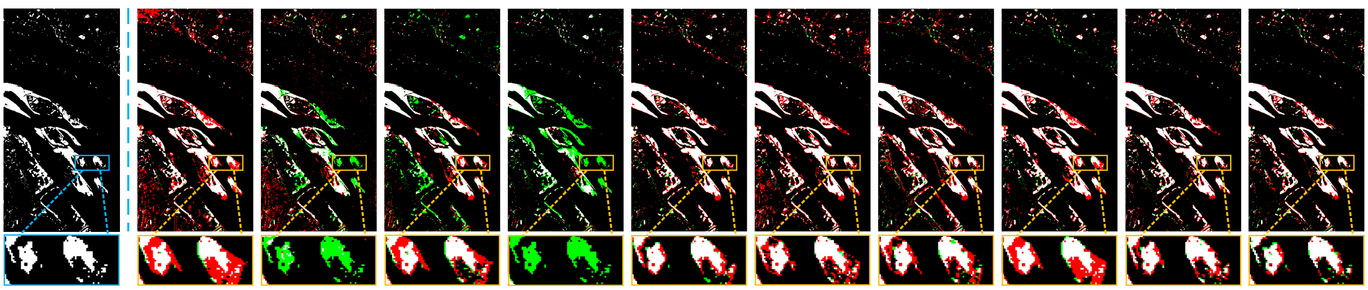
• UCH and CH are unchanged and changed samples, respectively.

C. Experimental Results

1) Experimental Results on the River Dataset: Table III and Fig. 7 provide the quantitative and qualitative evaluation results of all CD methods (including model-driven, data-driven, and joint-driven) on River dataset. In Table III, it can be seen that CVA, PCAKM, IR-MAD, and SVM, which are the traditional model-driven CD methods, achieve poor performance in terms of Precision and Recall metrics. For instance, CVA achieves 0.9627 in Recall but 0.5444 in Precision; SVM achieves 0.8652 in Precision metrics but only 0.4839 in Recall metrics. For data-driven methods, the CSANet and DECNet methods are able to achieve a better balance than the above-mentioned model-driven methods in terms of under-detection and over-detection, and their OA metrics achieved 0.9551 and 0.9569, respectively. However, it can be also observed that MSDFFN and ReCNN methods are lower than IR-MAD and SVM in terms of OA metrics, which can be attributed to the fact that the labeled samples with a ratio of 3% are difficult to allow MSDFFN and ReCNN to be adequately trained and learned. The joint-driven approaches (i.e., MSCSCNet and LRSRNet) achieve better results in terms of OA, Kappa, and, F1 metrics, which means that the joint-driven approaches can effectively integrate the advantages of model-driven and data-driven approaches to improve the HSI-CD performance. In contrast, the proposed LRSRNet method achieves the highest

TABLE III
QUALITY METRICS OBTAINED BY VARIOUS HSI-CD METHODS ON RIVER, FARMLAND, AND SANTA BARBARA DATASETS. NOTE THAT BOLD INDICATES THE HIGHEST ACCURACY.

Dataset		Different CD methods on the Bi-HSIs datasets.									
		Model-driven				**Data-driven**				**Joint-driven**	
		CVA	**PCAKM**	**IR-MAD**	**SVM**	**CSANet**	**MSDFFN**	**ReCNN**	**DECNet**	**MSCSCNet**	**LRSRNet**
River	OA	0.9267	0.9360	0.9490	0.9486	0.9551	0.9403	0.9424	0.9569	0.9625	**0.9648**
	Kappa	0.6575	0.6147	0.6820	0.5955	0.7577	0.7052	0.7011	0.7663	0.7954	**0.8033**
	Precision	0.5444	0.6198	0.7022	**0.8652**	0.6763	0.5972	0.6143	0.6857	0.7122	0.7314
	Recall	**0.9627**	0.6829	0.7178	0.4839	0.9271	0.9616	0.9055	0.9308	0.9545	0.9397
	F1	0.6955	0.6498	0.7099	0.6207	0.7821	0.7368	0.7320	0.7897	0.8157	**0.8225**
Farmland	OA	0.9523	0.9510	0.7765	0.8270	0.9715	0.9627	0.9692	0.9733	0.9732	**0.9744**
	Kappa	0.8855	0.8782	0.3567	0.5765	0.9307	0.9107	0.9251	0.9352	0.9357	**0.9380**
	Precision	0.9024	0.9514	0.7443	0.7074	0.9515	0.9150	0.9500	0.9525	0.9347	**0.9531**
	Recall	0.9370	0.8758	0.3500	0.6881	0.9501	0.9606	0.9436	0.9557	**0.9757**	0.9590
	F1	0.9194	0.9121	0.4761	0.6977	0.9508	0.9372	0.9468	0.9541	0.9548	**0.9560**
Santa Barbara	OA	0.8325	0.7755	0.7408	0.7221	0.9563	0.9629	0.8425	0.9329	0.9760	**0.9788**
	Kappa	0.6519	0.5458	0.3919	0.3815	0.9070	0.9229	0.6655	0.8612	0.9496	**0.9554**
	Precision	0.7741	0.6782	0.9352	0.7194	**0.9895**	0.9346	0.8246	0.8875	0.9769	0.9876
	Recall	0.8107	0.8166	0.3664	0.4809	0.8985	**0.9739**	0.7616	0.9498	0.9617	0.9582
	F1	0.7920	0.7410	0.5265	0.5765	0.9418	0.9538	0.7918	0.9176	0.9693	**0.9727**



(a) (b) (c) (d) (e) (f) (g) (h) (i) (j) (k)

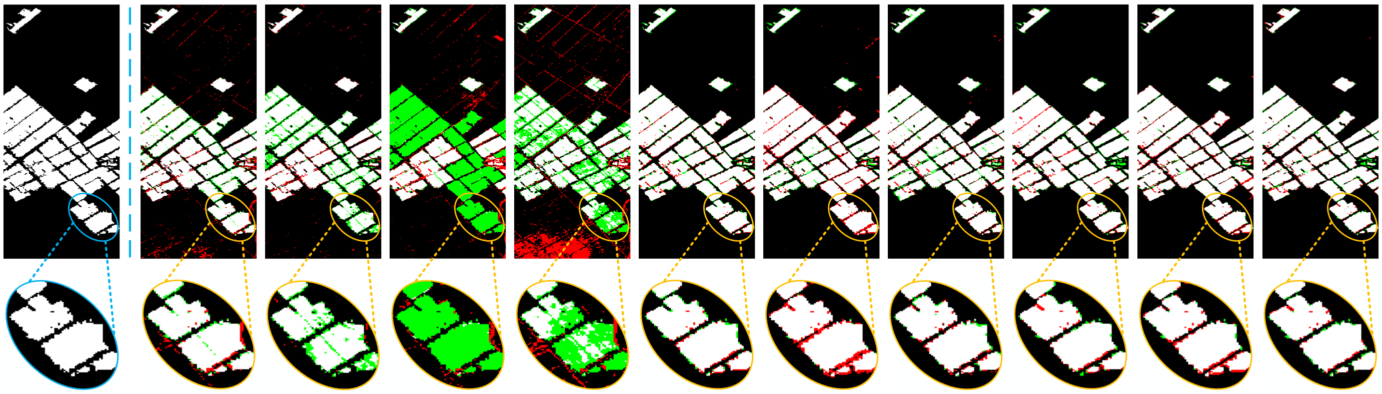
Fig. 7. CD results of various methods on the River dataset. (a) Groundtruth (GT). (b) CVA. (c) PCAKM. (d) IR-MAD. (e) SVM. (f) CSANet. (g) MSDFFN. (h) ReCNN. (i) MSCSCNet. (j) LRSRNet. Note that ■ indicates FP (i.e., over-detection) and ■ represents FN (i.e., under-detection).

accuracy with limited training samples compared with the MSCSCNet.

Fig. 7 presents the detection results of the GT and each CD method, and the over-detection (FP) and under-detection (FN) are marked with red and green, respectively. CVA maintains a high FP rate, which leads to a high recall rate. SVM has a high FN rate due to its strict judgment of changed features. The PCAKM and IR-MAD methods are similarly varying degrees of over-detection and under-detection. DL-based data-driven methods methods (i.e., the CSANet, MSDFFN, ReCNN, and DECNet methods) exhibit high FP rate and FN rate due to their sensitivity to the changed features between Bi-HSIs. The detection results of the joint-driven method (e.g., MSCSCNet) are significantly improved compared with the model- and data-driven methods. Compared with all CD methods, the FP rate and FN rate in the detection results of the proposed LRSRNet method in this paper are relatively lower than those of competing methods and closer to GT. This means that LRSRNet can achieve better detection performance compared to other CD methods.

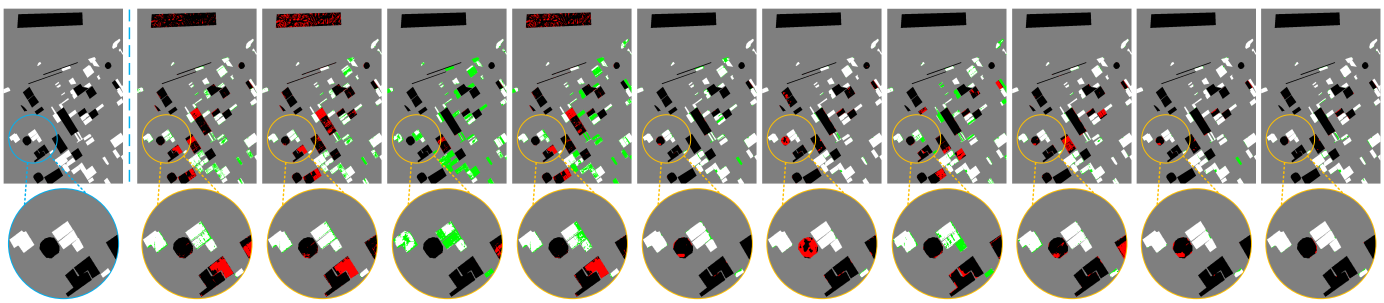
2) Experimental Results on the Farmland Dataset: Table III and Fig. 8 present the results of the performance test for each CD method on the Farmland dataset. This dataset contains information on farmland cover changes and noise subject to sensor imaging. It can be observed from Table III that model-driven methods are lower than data-driven methods in terms of five performance evaluation metrics. For instance, the OA metric of CVA is 0.9523, while the OA metrics of CSANet, MSDFFN, ReCNN, and DECNet are 0.9715, 0.9627, 0.9692, and 0.9733, respectively. This suggests that deep network architectures through RNNs and CNNs are able to efficiently extract the information about the differences between Bi-HSIs, which can enhance the CD result effect. The MSCSCNet adopting a joint driving strategy achieves better performance than model and data driven methods by designing multi-scale sparse coefficient features on the Farmland dataset. the proposed LRSRNet method, as a new joint-driven approach, adopts the strategy of intrinsic feature extraction and interference information separation to capture the real variation features, and thus achieving the highest accuracies in OA, Kappa, Precision, and F1, with 0.9744, 0.9344 and 0.9333, respectively. This directly points out the superior performance of the LRSRNet method using low-rank and sparse representation deep features for HSI-CD.

In addition, a visualization of the detection results of each CD method is given in Fig. 8. It can be clearly seen from elliptical window that the model-driven methods exhibit different degrees of over-detection and under-detection, especially the tendency to misclassify the presence of noisy areas as change areas. The data-driven and joint-driven (i.e., MSCSCNet) methods achieve promising results in terms of FP and FN rates, but are prone to over-detection and under-detection in ridge



(a) (b) (c) (d) (e) (f) (g) (h) (i) (j) (k)

Fig. 8. CD results of various methods on the Farmland dataset. (a) GT. (b) CVA. (c) PCAKM. (d) IR-MAD. (e) SVM. (f) CSANet. (g) MSDFFN. (h) ReCNN. (i) MSCSCNet. (j) LRSRNet. Note that ■ indicates FP (i.e., over-detection) and ■ represents FN (i.e., under-detection).



(a) (b) (c) (d) (e) (f) (g) (h) (i) (j) (k)

Fig. 9. CD results of various methods on the Santa Barbara dataset. (a) GT. (b) CVA. (c) PCAKM. (d) IR-MAD. (e) SVM. (f) CSANet. (g) MSDFFN. (h) ReCNN. (i) MSCSCNet. (j) LRSRNet. Note that ■ indicates FP (i.e., over-detection) and ■ represents FN (i.e., under-detection).

areas. Comparing the model- and data-driven methods, the LRSRNet achieved more promising detection results. The reasons for this are twofold: 1) the use of deep low-rank features to characterize the change information, which enhances the discriminative properties of changed and unchanged pixels; 2) the interference of noise information avoided by the separation and suppression of sparse features.

3) Experimental Results on the Santa Barbara Dataset: To further evaluate the performance of all CD methods, Bi-HSIs of complex scenarios are used for experiments. Table III and Fig. 9 show the HSI-CD results of model-, data-, and joint-driven methods on the Santa Barbara dataset. It can be seen from Table III that the LRSRNet method achieved the highest accuracy in OA, Kappa, and F1 metrics, which means that the LRSRNet still maintains a significant advantage in the HSI-CD task for complex scenes. However, the LRSRNet achieves sub-optimal accuracy in Precision compared to CSANet methods, which is attributed to the imbalance between CSANet methods in over-detection and under-detection, resulting in high Precision metric. Moreover, it can be clearly observed from Fig. 9 (see circular area) that the LRSRNet method achieves significant detection results in terms of over-detection and under-detection compared to all the competing methods (including model-, data-, and joint-driven) in the comparative experimental schedule. This further demonstrates the effectiveness and utility of LRSRNet in the HSI-CD task.

TABLE IV
MCNEMAR'S TEST BETWEEN THE PROPOSED METHOD AND OTHER CHANGE DETECTION APPROACHES.

Methods	The significance statistic $\|Z\|$		
	River	Farmland	Santa Barbara
LRSRNet vs CVA	47.969	29.458	132.17
LRSRNet vs PCAKM	34.606	31.713	157.97
LRSRNet vs IR-MAD	21.802	107.51	173.96
LRSRNet vs SVM	20.335	90.098	186.51
LRSRNet vs CSANet	15.460	10.753	39.129
LRSRNet vs MSDFFN	34.785	19.608	29.216
LRSRNet vs ReCNN	32.379	13.051	124.33
LRSRNet vs DECNet	13.014	8.321	63.156
LRSRNet vs MSCSCNet	4.057	8.003	5.992

• $|Z| \geq 1.96$ is significant, the larger the value, the more significant.

D. Statistical Significance Analysis of the Results

In this section, the McNemar's test [69], [70] is employed to evaluate the statistical significance of the differences between the different HSI-CD methods. The standardized normal test statistic Z can be described as follows:

$$Z = \frac{\mathcal{A}_{12} - \mathcal{A}_{21}}{\sqrt{\mathcal{A}_{12} + \mathcal{A}_{21}}} \tag{37}$$

where $\mathcal{A}_{ij}$ represents the number of samples that are correctly detected by detector i but incorrectly detected by detector j. Z represents the pairwise statistical significance of the difference in change detection between the i-th and j-th detectors. Note that the test statistic $|Z| \geq 1.96$ indicates that the difference in change detection accuracy between the i-th and j-th detectors

TABLE V
ABLATION STUDY FOR THE PROPOSED LRSRNET METHOD ON THE RIVER DATASET. THE BOLD TYPE IS USED FOR THE OPTIMAL INDICATOR, AND THE UNDERLINE IS USED FOR THE SUB-OPTIMAL INDICATOR.

Ablation stage		LRSD			RALM	Quantitative metrics of performance contribution						
		LREM	SEM	HRM		OA	Kappa	Precision	Recall	F1	Param.	FLOPs
LRSNet	With	✓	✓	✓	✓	**0.9629**	**0.7934**	**0.7215**	**0.9328**	**0.8137**	4.070M	359.302M
	Without	✗				0.9493	0.7304	0.6483	0.9117	0.7578	3.120M	265.040M
			✗			0.9469	0.7201	0.6363	0.9095	0.7487	3.120M	265.040M
				✗		0.9480	0.7252	0.6412	0.9128	0.7533	3.120M	264.938M
					✗	0.9532	0.7490	0.6668	0.9231	0.7743	3.600M	314.245M
	Baseline	✗	✗	✗	✗	0.9243	0.6214	0.5410	0.8514	0.6616	747.8K	29.302M
Ablation stage		LRSD	RALM			Quantitative metrics of performance contribution						
			RLM	SAM	CAM	OA	Kappa	Precision	Recall	F1	Param.	FLOPs
LRSRNet	With	✓	✓	✓	✓	**0.9629**	**0.7934**	**0.7215**	**0.9328**	**0.8137**	4.070M	359.302M
	Without	✗				0.9444	0.7107	0.6229	0.9130	0.7405	3.599M	314.180M
			✗			0.9539	0.7519	0.6700	0.9245	0.7769	4.042M	358.416M
				✗		0.9544	0.7524	0.6753	0.9151	0.7771	4.062M	359.108M
					✗	0.9575	0.7683	0.6901	0.9275	0.7914	4.048M	358.545M
	Baseline	✗	✗	✗	✗	0.9243	0.6214	0.5410	0.8514	0.6616	747.8K	29.302M

• RLM, SAM, and CAM stand for residual learning mechanism, spectral attention mechanism, and coordinate attention mechanism, respectively.

is considered statistically significant at the 5% significance level [69]. The statistical results of the McNemar's test of the LRSRNet and competitive methods are given in Table IV, which demonstrates that the proposed LRSRNet method is superior to model-driven (i.e., CVA, PCAKM, IR-MAD, and SVM), data-driven (i.e., CSANet, MSDFFN, ReCNN, and DECNet), and joint-driven (i.e., MSCSCNet) methods on the River, Farmland, and Santa Barbara. This means that LRSRNet designed with deep unfolding architecture of LRSR is able to improve the change detection accuracy of Bi-HSIs due to the statistically significant test statistics. Meanwhile, this further proves the effectiveness of the proposed LRSRNet method in the HSI-CD task.

E. Ablation Analysis

This section conducts ablation experiments on the proposed LRSRNet method on the River dataset to analyze the performance gains of core components. In fact, the LRSRNet method contains two sub-networks, i.e., LRSnet and CDnet. The former aims to capture the deep semantic features of change information from Bi-HSIs, while the latter focuses on the decision of pixel attributes (changed or unchanged). In this experiment, the important modules of LRSD and RALM in LRSnet are ablated one by one to analyze the performance contribution, and CDnet is used as a benchmark model to provide a standard reference. Note that LRSD includes the LREM, SEM, and HRM modules, while RALM includes the RLM, SAM, and CAM modules.

Table V presents the quantitative results of the above ablation experiments on the River dataset. It can be clearly observed that when the LREM, SEM, and HRM modules in LRSD are removed from LRSRNet respectively, its HSI-CD performance degrades significantly. It is worth mentioning that the performance degradation caused by the ablation of the LRSD module is greater than that of the ablation of the RALM module. This means that the three modules in LRSD bring more performance gain to LRSRNet than the RALM module, and also confirms the effectiveness of LRSR-based deep unfolding network in the HSI-CD task. Furthermore, the three modules in RALM are subjected to detailed ablation experiments, and it can be seen that the performance of LRSRNet also degrades as the RLM, SAM, and CAM modules are removed, respectively. This shows that the design of RLM, SAM, and CAM modules in RALM can effectively promote the accuracy of the LRSRNet method in change detection task.

F. Discussion

1) Discussion on the Number of Residual Layers: To investigate the effect of the number of residual layers (i.e., L_r in Section II-B) in RALM on the LRSRNet method in terms of performance and efficiency, we analyzed the detection performance of LRSRNet on three real Bi-HSIs by setting different values for L_r. As shown in Fig. 10, the orange and green rectangular heights in Figs. 10(a)-(c) represent the OA and Kappa metrics, respectively, and the curves are the F1 metrics. Fig.10(d) depicts the OA, Kappa, Param and FLOPs metrics for L_r at different values. It can be seen from Fig. 10 that the OA, Kappa, and F1 metrics vary significantly with the variation of L_r from 1 to 5 and from the overall analysis $L_r = 2$ is a node of turnaround in performance. In addition, as shown in Fig. 10(d), the values of $L_r = 1$ to $L_r = 4$ have a small impact on the detection performance of the LRSRNet, but the number of parameters and FLOPs increase noticeably. Therefore, considering the balance of detection performance over efficiency, $L_r = 2$ is set as the default parameter of the LRSRNet method.

2) Discussion on the Patch Size: The characterization of pixel changes between Bi-HSIs is often represented using a joint representation of the central pixel and its neighborhood information, and thus the patch size depends on the amount of neighborhood information. As shown in Fig. 11, we explored the effect of patch size on the performance of the LRSRNet method in the interval [5×5, 7×7, 9×9, 11×11, 13×13]. It can be observed that as the patch size is taken from 5×5 to 13×13, the LRSRNet method shows a trend of degradation in detection performance on the River, Farmland, and Santa

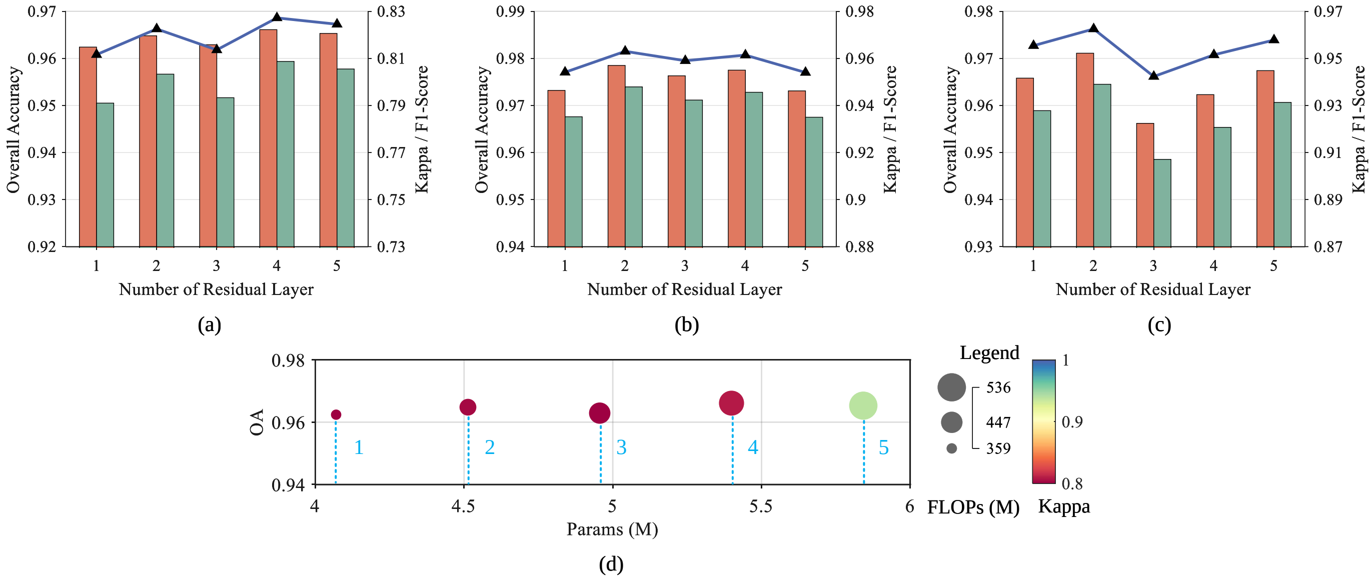


Fig. 10. CD accuracy achieved by the proposed LRSRNet method with different numbers of residual layers on three datasets. (a) River dataset. (b) Farmland dataset. (c) Santa Barbara dataset. (d) Network parameters and FLOPs on the (a).

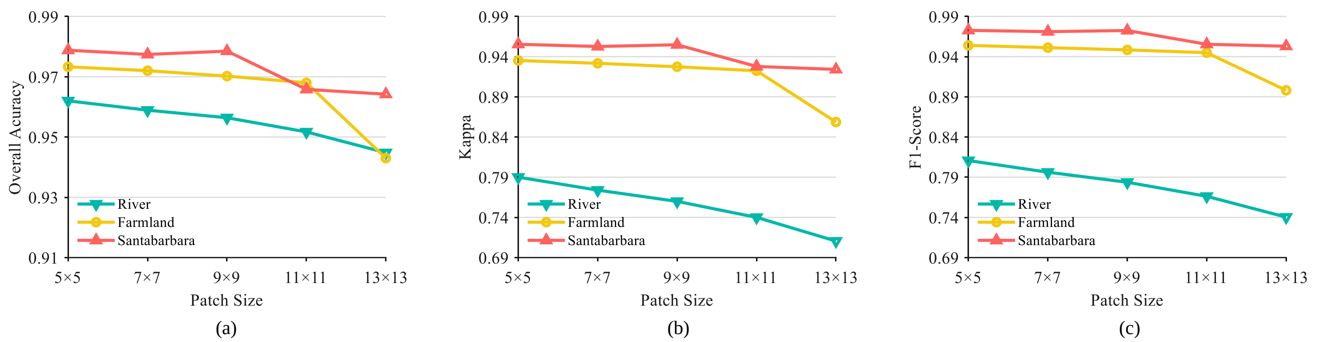


Fig. 11. Effects of different patch sizes on OA, Kappa, and F1-Score for the proposed LRSRNet method on River, Farmland, and Santa Barbara datasets. (a) OA. (b) Kappa. (c) F1-Score.

Barbara datasets. The reason is that the capture of Bi-HSIs is affected by the acquisition time and imaging sensors, and there are pseudo-change features (e.g., weather) in the larger neighborhood space, thus appropriate neighborhood spectral-spatial information is beneficial for identifying the real change region of Bi-HSIs.

TABLE VI
EFFECT OF STAGE NUMBER K ON THE CHANGED DETECTION PERFORMANCE ON THE FARMLAND DATASET.

Stages (K)	CD contribution with LRSD iteration					
	OA	**Kappa**	**Precision**	**Recall**	**F1**	**Params**
K=1	0.9621	0.9071	0.9481	0.9196	0.9336	1.58M
K=2	0.9741	0.9368	0.9663	0.9438	0.9549	2.41M
K=3	0.9742	0.9370	0.9637	0.9467	0.9551	3.24M
K=4	**0.9783**	**0.9474**	**0.9604**	**0.9650**	**0.9627**	4.07M
K=5	0.9778	0.9462	0.9581	0.9656	0.9619	4.90M
K=6	0.9742	0.9375	0.9548	0.9564	0.9556	5.73M
K=7	0.9736	0.9363	0.9465	0.9636	0.9550	6.56M

3) Discussion on LRSD Stages: We analyze the performance impact of the important parameter K in LRSnet on the proposed LRSRNet method on the Farmland dataset. We emphasize that the low-rank and sparse components of HSI are extracted more adequately as the value of K increases, but there is an impact on the execution efficiency. Table VI reports

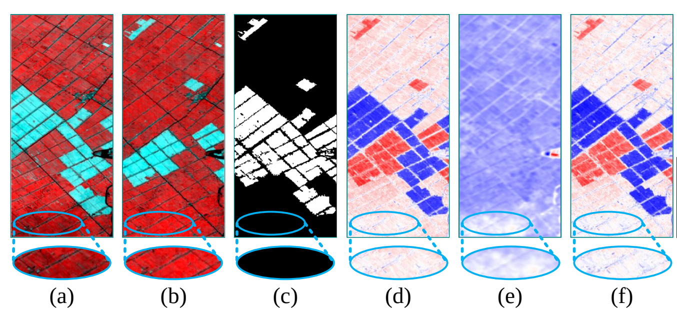


Fig. 12. Visualization of low-rank and sparse features from the LRSD on Farmland dataset. (a) and (b) are T_1 and T_2 HSI false color, respectively. (c) GT. (d)–(f) are first principal component of difference of Bi-HSIs for raw HSI, sparse feature, and low-rank feature, respectively.

the performance metrics of the LRSRNet for different values of K. Focusing on OA and Params, it can be observed that as the value of K increases, the OA metrics show an increasing and then decreasing trend, and Params shows a sharp increase. Therefore, according to the optimal performance principle, $K = 4$ is adopted as the optimal parameter of LRSnet to LRSRNet.

In addition, we provide quantitative and visualization results for the LRSRNet with and without LRSD process. In Table VII, the LRSRNet method with LRSD obtains better detection results than the LRSRNet method without LRSD. This implies that the low-rank features extracted by the LRSD process can facilitate the identification of change information between Bi-HSIs. Moreover, Fig. 12 presents the results of the visualization of the first principal component of difference of Bi-HSIs for low-rank and sparse features achieved by the LRSRNet, it can be observed from Fig. 12(d)–(f) that noise from raw HSIs can be effectively suppressed to enhance the saliency of changed and unchanged pixels (see elliptical region).

TABLE VII
CHANGE RESULTS OF THE PROPOSED METHOD WITH OR WITHOUT LRSD ON THE FARMLAND DATASET.

Component	CD contribution with or without LRSD				
	OA	**Kappa**	**Precision**	**Recall**	**F1**
without	0.9529	0.7602	0.9394	0.6755	0.7859
with	**0.9783**	**0.9474**	**0.9604**	**0.9650**	**0.9627**

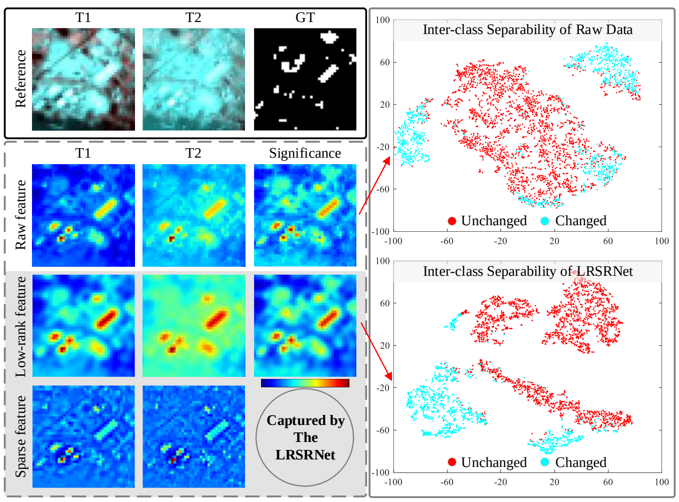


Fig. 13. Feature visualization of the proposed LRSRNet.

4) Discussion on Feature Visualization: To prove the effectiveness of the proposed LRSRNet method, Fig. 13 provides visual results for LRSRNet. The second row of the left area presents the significance of the raw data T1 and T2 images structure and texture details as well as the changed pixels of the T1 and T2 images. In contrast, the low-rank features captured by the LRSRNet method can effectively perceive the intrinsic information of the change in ground coverage between T1 and T2 images, and enhance the details of the changed pixels (obtained by LREM). The sparse features captured by the LRSRNet method can obtain the interference information of the changed pixels and suppress the noise expression (obtained by SEM). Furthermore, we visualize the raw data and the discriminative features learned by the LRSRNet method in the right panel (i.e., using t-SNE), and it can be clearly seen that the discriminative features obtained by the LRSRNet method have significant inter-class separability compared to the raw data on the River dataset. Therefore, this undoubtedly proves that the proposed LRSRNet method has the ability to guide discriminative feature extraction using low-rank and sparse unfolding networks and its effectiveness in the HSI-CD task.

5) Discussion on LRSnet Stages: LRSnet is an important component of the proposed LRSRNet method. LRSnet consists of three modules, i.e., LREM, SEM, and LRM, where t_l, t_s, and t_h determine the network depth of each of the three modules. In this section, we investigate the effect of t_l, t_s, and t_h on the detection performance of LRSRNet. The value intervals of the t_l, t_s, and t_h parameters are set to [1, 3, 6, 9, 12]. It can be seen from Fig. 14 that the LRSRNet achieves poor detection performance when the above parameter is set to 1, while the detection performance of the LRSRNet has a small fluctuation when the above parameter is gradually set to 3,6,9,12. Moreover, the larger the values of the t_l, t_s, and t_h parameters, the larger the number of parameters of LRSRNet. Therefore, the t_l, t_s, and t_h parameters are all set to 3 based on a consideration of performance and execution efficiency.

6) Discussion on the Loss Function: The loss function of this work consists of two parts, on the one hand, it considers the data fidelity of the reconstructed HSIs output by HRM, and on the other hand, it takes into account the accuracy of the CDnet for the detection of the change information between Bi-HSIs. The data fidelity and detection accuracy are constrained using $\mathcal{L}_1$ and $\mathcal{L}_{CE}$, respectively. Table VIII reports the performance metrics of LRSRNet under $\mathcal{L}_{CE}$ working alone and its performance metrics under $\mathcal{L}_{CE}$ and $\mathcal{L}_1$ working jointly on the River, Farmland, and Santa Barbara datasets. It can be observed that the joint $\mathcal{L}_{CE}$ and $\mathcal{L}_1$ loss function is able to drive LRSRNet to achieve promising HSI-CD performance. In addition, we investigated the trade-off parameter λ in the above loss function in Fig. 15, and it can be seen that LRSRNet achieves poor detection performance when $\lambda = 0$, which indicates that the use of $\mathcal{L}_{CE}$ alone does not enable LRSRNet to learn effectively. The reason for this is that the deep network does not take into account the fidelity of the HSI reconstruction during the training stage, which leads to the imprecise extraction of the low-rank components.

TABLE VIII
CD PERFORMANCE OF THE PROPOSED LRSRNET METHOD WITH VARIOUS LOSS FUNCTIONS ON RIVER, FARMLAND, AND SANTA BARBARA DATASET, RESPECTIVELY.

Metrics	River		Farmland		Santa Barbara	
	$\mathcal{L}_{CE}$	$+\mathcal{L}_1$	$\mathcal{L}_{CE}$	$+\mathcal{L}_1$	$\mathcal{L}_{CE}$	$+\mathcal{L}_1$
OA	0.9427	**0.9667**	0.8820	**0.9779**	0.8555	**0.9774**
Kappa	0.7114	**0.8105**	0.6796	**0.9464**	0.6775	**0.9526**
Precision	0.6098	**0.7484**	0.9430	**0.9563**	0.9549	**0.9767**
Recall	0.9471	**0.9285**	0.6171	**0.9679**	0.6425	**0.9656**
F1	0.7419	**0.8288**	0.7522	**0.9621**	0.7776	**0.9711**

7) Discussion on Computational Cost: This section presents the time consumption of all CD methods in Table IX on the River dataset. Although model-driven algorithms, such as CVA, PCAKM, and IR-MAD, achieve poor detection performance in terms of balance between under-detection and over-detection, these methods have shown high execution efficiency. The CD detection performance of the data-driven method is more advantageous than that of the model-driven method, but it is limited by the number of parameters and FLOPs of the deep network, which takes up more time. The MSCSCNet and LRSRNet belong to the deep network

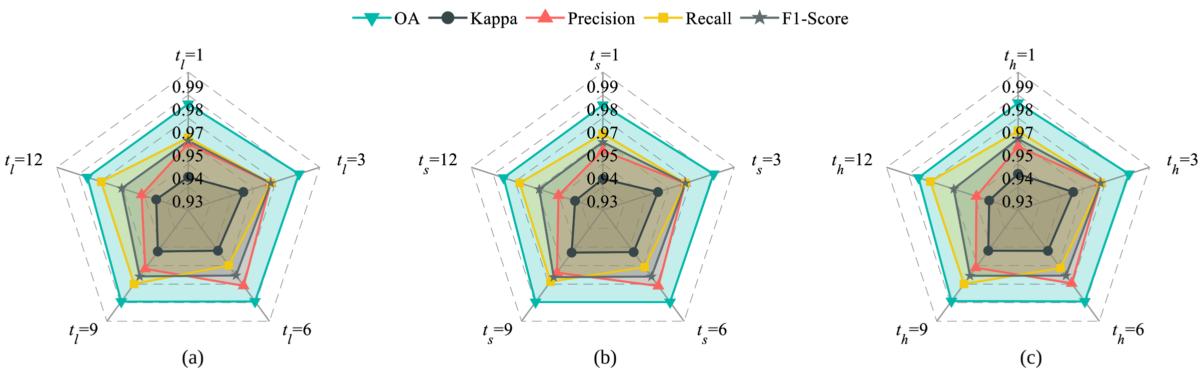


Fig. 14. CD results obtained by the proposed LRSRNet with different layers within sparse learning, low-rank learning, and image recovery stage on the Farmland dataset. (a) Low-rank learning stage t_l. (b) Sparse learning stage t_s. (c) Image recovery stage t_h.

TABLE IX
COMPUTATION COST OF VARIOUS METHODS ON RIVER DATASET.

Method	Model-driven				Data-driven				Joint-driven	
	CVA	PCAKM	IR-MAD	SVM	CSANet	MSDFFN	ReCNN	DECNet	MSCSCNet	LRSRNet
Params	/	/	/	/	2.4679M	88.771M	878.59K	16.006M	32.224M	4.2992M
FLOPs	/	/	/	/	293.52M	581.97M	5.9034M	520.56M	81.895G	359.30M
Running Time (s)	0.0827	0.1423	0.6997	4.2468	5.7110	6.0349	1.7834	5.5102	150.25	5.9197

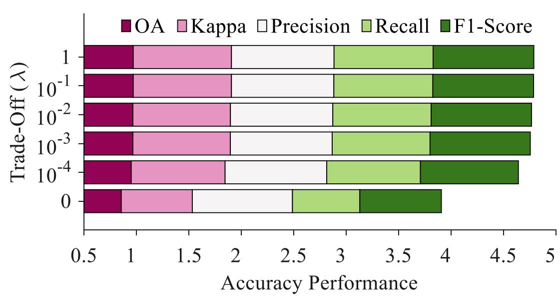


Fig. 15. CD results achieved by the proposed LRSRNet with different trade-off parameters in training stage on the Santa Barbara dataset.

architecture jointly driven by the model and data, aiming to enhance the transparency of the network through model-guided architecture design. The MSCSCNet method achieves better CD accuracy than the above model-driven and data-driven methods. In contrast, the proposed LRSRNet method not only maintains its advantage in CD accuracy, but also achieves acceptable performance in terms of number of parameters, FLOPs and execution time compared to the advanced the MSCSCNet.

V. CONCLUSIONS

In this paper, we propose a new low-rank and sparse representation-based deep unfolding network (LRSRNet) for change detection with Bi-HSIs, which has a distinct feature mining mechanism and a transparent architecture design. We have designed LRSnet and CDnet sub-networks in the proposed LRSRNet architecture. LRSnet is utilized to extract the valuable low-rank information of HSI whilst suppressing the noise-containing sparse components, and CDnet is employed to discriminate the change information on the low-rank features learnt by LRSnet, which fully defines the process and mechanism of feature mining in the HSI-CD task. In addition, we construct a transparent deep network structure by applying the deep unfolding network paradigm to the low-rank and sparse decomposition process to gradually learn and optimize the low-rank and sparse components in HSI. Experiments on three real datasets have demonstrated the superior performance of our LRSRNet method over nine model-driven, data-driven, and model-data-joint-driven HSI-CD algorithms.

In the future, we will work on two research lines: 1) integrating feature extraction and change detection into a uniform deep unfolding network to further improve the flexibility of deep network in HSI-CD tasks; and 2) to deeply combine the LRSRNet method with the few-shot learning so as to address the challenges of HSI in large-scale and long-time-series cross-modal heterogeneous change detection applications [71].

ACKNOWLEDGMENT

The authors would like to thank the Editor-in-Chief, the anonymous Associate Editor, and the reviewers for their insightful comments and suggestions, which significantly improved the quality and presentation of this article.

REFERENCES

[1] S.-E. Qian, "Hyperspectral satellites, evolution, and development history," *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.*, vol. 14, p. 7032–7056, 2021.

[2] Y. Su, J. Chen, L. Gao, A. Plaza, M. Jiang, X. Xu, X. Sun, and P. Li, "Acgt-net: Adaptive cuckoo refinement-based graph transfer network for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 61, p. 1–14, Aug. 2023.

[3] Y. Su, Z. Zhu, L. Gao, A. Plaza, P. Li, X. Sun, and X. Xu, "Daan: A deep autoencoder-based augmented network for blind multilinear hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 62, p. 1–15, Mar. 2024.

[4] A. S. Sahadevan, "Extraction of spatial-spectral homogeneous patches and fractional abundances for field-scale agriculture monitoring using airborne hyperspectral images," *Comput. Electron. Agric.*, vol. 188, p. 106325, Sep. 2021.

[5] D. Marinelli, F. Bovolo, and L. Bruzzone, "A novel change detection method for multitemporal hyperspectral images based on binary hyperspectral change vectors," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 7, p. 4913–4928, Jul. 2019.

[6] L. Zhang and J. Xia, "Flood detection using multiple chinese satellite datasets during 2020 china summer floods," *Remote Sens.*, vol. 14, no. 1, p. 51, Dec. 2021.

[7] S. Liu, D. Marinelli, L. Bruzzone, and F. Bovolo, "A review of change detection in multitemporal hyperspectral images: Current techniques, applications, and challenges," *IEEE Geosci. Remote Sens. Mag.*, vol. 7, no. 2, p. 140–158, Jun. 2019.

[8] M. Hasanlou and S. T. Seydi, "Hyperspectral change detection: an experimental comparative study," *Int. J. Remote Sens.*, vol. 39, no. 20, p. 7029–7083, Apr. 2018.

[9] F. Bovolo and L. Bruzzone, "A theoretical framework for unsupervised change detection based on change vector analysis in the polar domain," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 1, p. 218–236, Jan. 2007.

[10] Z. Hou, W. Li, L. Li, R. Tao, and Q. Du, "Hyperspectral change detection based on multiple morphological profiles," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, p. 1–12, Jul. 2022.

[11] T. Celik, "Unsupervised change detection in satellite images using principal component analysis and k-means clustering," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 4, p. 772–776, Oct. 2009.

[12] A. A. Nielsen, "The regularized iteratively reweighted mad method for change detection in multi- and hyperspectral data," *IEEE Trans. Image Process.*, vol. 16, no. 2, p. 463–478, Feb. 2007.

[13] C. Wu, B. Du, and L. Zhang, "Slow feature analysis for change detection in multispectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 5, p. 2858–2874, May 2014.

[14] T. Habib, J. Inglada, G. Mercier, and J. Chanussot, "Support vector reduction in svm algorithm for abrupt change detection in remote sensing," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 3, p. 606–610, Jul. 2009.

[15] D. Gevana, L. Camacho, A. Carandang, S. Camacho, and S. Im, "Land use characterization and change detection of a small mangrove area in banacon island, bohol, philippines using a maximum likelihood classification method," *For. Sci. Technol.*, vol. 11, no. 4, p. 197–205, Apr. 2015.

[16] T. Abeywickrama, M. A. Cheema, and D. Taniar, "k-nearest neighbors on road networks: A journey in experimentation and in-memory implementation," 2016.

[17] M. Hu, C. Wu, and L. Zhang, "Hypernet: Self-supervised hyperspectral spatial–spectral feature understanding network for hyperspectral change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, p. 1–17, Nov. 2022.

[18] F. Luo, T. Zhou, J. Liu, T. Guo, X. Gong, and J. Ren, "Multiscale diff-changed feature fusion network for hyperspectral image change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, p. 1–13, Jan. 2023.

[19] Y. Wang, D. Hong, J. Sha, L. Gao, L. Liu, Y. Zhang, and X. Rong, "Spectral–spatial–temporal transformers for hyperspectral image change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, p. 1–14, Aug. 2022.

[20] M. Gong, F. Jiang, A. K. Qin, T. Liu, T. Zhan, D. Lu, H. Zheng, and M. Zhang, "A spectral and spatial attention network for change detection in hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, p. 1–14, Dec. 2022.

[21] L. Mou, L. Bruzzone, and X. X. Zhu, "Learning spectral-spatial-temporal features via a recurrent convolutional neural network for change detection in multispectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 2, p. 924–935, Feb. 2019.

[22] 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, p. 9976–9992, Dec. 2019.

[23] Y. Li, J. Ren, Y. Yan, Q. Liu, P. Ma, A. Petrovski, and H. Sun, "Cbanet: An end-to-end cross-band 2-d attention network for hyperspectral change detection in remote sensing," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, p. 1–11, 2023.

[24] H. Chen, C. Wu, B. Du, L. Zhang, and L. Wang, "Change detection in multisource vhr images via deep siamese convolutional multiple-layers recurrent neural network," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 4, p. 2848–2864, Apr. 2020.

[25] C. Shi, Z. Zhang, W. Zhang, C. Zhang, and Q. Xu, "Learning multiscale temporal–spatial–spectral features via a multipath convolutional lstm neural network for change detection with hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, p. 1–16, May 2022.

[26] A. Graves, *Long Short-Term Memory*, Jan. 2012, p. 37–45.

[27] Q. Wang, Z. Yuan, Q. Du, and X. Li, "Getnet: A general end-to-end 2-d cnn framework for hyperspectral image change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 1, p. 3–13, Jan. 2019.

[28] X. Ou, L. Liu, B. Tu, L. Qing, G. Zhang, and Z. Liang, "Cbw-mssanet: A cnn framework with compact band weighting and multiscale spatial attention for hyperspectral image change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, p. 1–18, Mar. 2023.

[29] J. Qu, W. Dong, Y. Yang, T. Zhang, Y. Li, and Q. Du, "Cycle-refined multidecision joint alignment network for unsupervised domain adaptive hyperspectral change detection," *IEEE Trans. Neural Netw. Learn. Syst.*, p. 1–14, Jan. 2024.

[30] X. Wang, K. Zhao, X. Zhao, and S. Li, "Tritf: A triplet transformer framework based on parents and brother attention for hyperspectral image change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, p. 1–13, Mar. 2023.

[31] Y. Xiao, Q. Yuan, K. Jiang, J. He, C.-W. Lin, and L. Zhang, "Ttst: A top-k token selective transformer for remote sensing image super-resolution," *IEEE Trans. Image Process.*, vol. 33, p. 738–752, Jan. 2024.

[32] A. Shafique, S. T. Seydi, T. Alipour-Fard, G. Cao, and D. Yang, "Ssvit-hcd: A spatial–spectral convolutional vision transformer for hyperspectral change detection," *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.*, vol. 16, p. 6487–6504, Mar. 2023.

[33] C. Zhou, Q. Shi, D. He, B. Tu, H. Li, and A. Plaza, "Spectral-spatial sequence characteristics-based convolutional transformer for hyperspectral change detection," *CAAI Trans. Intell. Tech.*, vol. 8, no. 4, p. 1237–1257, May 2023.

[34] A. Hohl, I. Obadic, M.-A. Fernandez-Torres, H. Najjar, D. A. B. Oliveira, Z. Akata, A. Dengel, and X. X. Zhu, "Opening the black box: A systematic review on explainable artificial intelligence in remote sensing," *IEEE Geosci. Remote Sens. Mag.*, vol. 12, no. 4, p. 261–304, Dec. 2024.

[35] J. Li, K. Zheng, L. Gao, Z. Han, Z. Li, and J. Chanussot, "Enhanced deep image prior for unsupervised hyperspectral image super-resolution," *IEEE Trans. Geosci. Remote Sens.*, vol. 63, p. 1–18, Jan. 2025.

[36] J. Li, K. Zheng, L. Gao, L. Ni, M. Huang, and J. Chanussot, "Model-informed multistage unsupervised network for hyperspectral image super-resolution," *IEEE Trans. Geosci. Remote Sens.*, vol. 62, p. 1–17, Apr. 2024.

[37] C. Li, B. Zhang, D. Hong, J. Yao, and J. Chanussot, "Lrr-net: An interpretable deep unfolding network for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, p. 1–12, May 2023.

[38] C. Li, B. Zhang, D. Hong, J. Yao, X. Jia, A. Plaza, and J. Chanussot, "Interpretable networks for hyperspectral anomaly detection: A deep unfolding solution," *IEEE Trans. Geosci. Remote Sens.*, vol. 62, p. 1–14, Jul. 2024.

[39] X. Nie, Z. Xue, H. Su, and J. Li, "Nsr-net: Representation model-inspired interpretable deep unfolding network for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 63, p. 1–16, Dec. 2025.

[40] Y. Xiao, Q. Yuan, K. Jiang, J. He, X. Jin, and L. Zhang, "Ediffsr: An efficient diffusion probabilistic model for remote sensing image super-resolution," *IEEE Trans. Geosci. Remote Sens.*, vol. 62, p. 1–14, Dec. 2024.

[41] K. Zhang, J. Yan, F. Zhang, C. Ge, W. Wan, J. Sun, and H. Zhang, "Spectral–spatial dual graph unfolding network for multispectral and hyperspectral image fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 62, p. 1–18, Feb. 2024.

[42] J. Zhao, S. Xiao, W. Dong, J. Qu, and Y. Li, "Dictionary learning-guided deep interpretable network for hyperspectral change detection," *IEEE Geosci. Remote Sens. Lett.*, vol. 20, p. 1–5, Aug. 2023.

[43] J. Qu, P. Yang, W. Dong, X. Zhang, and Y. Li, "A semi-supervised multiscale convolutional sparse coding-guided deep interpretable network for hyperspectral image change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 62, p. 1–14, Sep. 2024.

[44] J. Peng, W. Sun, H.-C. Li, W. Li, X. Meng, C. Ge, and Q. Du, "Low-rank and sparse representation for hyperspectral image processing: A review," *IEEE Geosci. Remote Sens. Mag.*, vol. 10, no. 1, p. 10–43, Mar. 2022.

[45] C. Cao, J. Yu, C. Zhou, K. Hu, F. Xiao, and X. Gao, "Hyperspectral image denoising via subspace-based nonlocal low-rank and sparse factorization," *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.*, vol. 12, no. 3, p. 973–988, Mar. 2019.

[46] A. Soltani-Farani, H. R. Rabiee, and S. A. Hosseini, "Spatial-aware dictionary learning for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 1, p. 527–541, Jan. 2015.

[47] T. Cheng and B. Wang, "Graph and total variation regularized low-rank representation for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 1, p. 391–406, Jan. 2020.

[48] Z. Chen and B. Wang, "Spectrally-spatially regularized low-rank and sparse decomposition: A novel method for change detection in multi-temporal hyperspectral images," *Remote Sens.*, vol. 9, no. 10, p. 1044, Oct. 2017.

[49] J. Wright, A. Ganesh, S. R. Rao, Y. Peng, and Y. Ma, "Robust principal component analysis: Exact recovery of corrupted low-rank matrices via convex optimization," in *Proc. Adv. Neural Inf. Process. Syst. (NIPS)*, 2009, pp. 2080–2088.

[50] Z. Zhou, X. Li, J. Wright, E. Candes, and Y. Ma, "Stable principal component pursuit," in *Proc. IEEE Int. Symp. Inf. Theor.*, Jun. 2010, pp. 1518–1522.

[51] 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, p. 1–5, Jun. 2022.

[52] 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, p. 2418–2422, Dec. 2017.

[53] H. Zhang, W. He, L. Zhang, H. Shen, and Q. Yuan, "Hyperspectral image restoration using low-rank matrix recovery," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 8, p. 4729–4743, Aug. 2014.

[54] J. Peng, W. Sun, H.-C. Li, W. Li, X. Meng, C. Ge, and Q. Du, "Low-rank and sparse representation for hyperspectral image processing: A review," *IEEE Geosci. Remote Sens. Mag.*, vol. 10, no. 1, p. 10–43, Mar. 2022.

[55] H. Li and Z. Lin, "Accelerated proximal gradient methods for nonconvex programming," in *Proc. Int. Conf. Neural Info. Process. Syst.*, ser. NIPS'15, Cambridge, MA, USA, 2015, p. 379–387.

[56] S. Ma, D. Goldfarb, and L. Chen, "Fixed point and bregman iterative methods for matrix rank minimization," *Math. Program.*, vol. 128, no. 1–2, p. 321–353, Sep. 2009.

[57] H. Cai, J. Liu, and W. Yin, "Learned robust pca: A scalable deep unfolding approach for high-dimensional outlier detection," *Adv. Neural Inf. Process. Syst.*, vol. 34, no. 16977–16989, p. 16977–16989, Oct. 2021.

[58] T. Zhang, L. Li, C. Igel, S. Oehmcke, F. Gieseke, and Z. Peng, "Lr-csnet: Low-rank deep unfolding network for image compressive sensing," in *2022 IEEE 8th International Conference on Computer and Communications (ICCC)*, Dec. 2022, p. 1951–1957.

[59] J. Zhang and B. Ghanem, "Ista-net: Interpretable optimization-inspired deep network for image compressive sensing," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2018, p. 1828–1837.

[60] J. Mifdal, M. Tomás-Cruz, A. Sebastianelli, B. Coll, and J. Duran, "Deep unfolding for hyper sharpening using a high-frequency injection module," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2023, p. 2106–2115.

[61] H. Wang, Y. Li, H. Zhang, D. Meng, and Y. Zheng, "Indudonet+: A deep unfolding dual domain network for metal artifact reduction in ct images," *Med. Image Anal.*, vol. 85, p. 102729, Apr. 2023.

[62] Q. Wang, B. Wu, P. Zhu, P. Li, W. Zuo, and Q. Hu, "Eca-net: Efficient channel attention for deep convolutional neural networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, p. 11531–11539.

[63] Q. Hou, D. Zhou, and J. Feng, "Coordinate attention for efficient mobile network design," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 13 708–13 717.

[64] L. Bruzzone and D. Prieto, "Automatic analysis of the difference image for unsupervised change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 3, p. 1171–1182, May 2000.

[65] R. Song, W. Ni, W. Cheng, and X. Wang, "Csanet: Cross-temporal interaction symmetric attention network for hyperspectral image change detection," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, p. 1–5, May 2022.

[66] F. Luo, T. Zhou, J. Liu, T. Guo, X. Gong, and J. Ren, "Multiscale diff-changed feature fusion network for hyperspectral image change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, p. 1–13, Jan. 2023.

[67] L. Mou, L. Bruzzone, and X. X. Zhu, "Learning spectral-spatial-temporal features via a recurrent convolutional neural network for change detection in multispectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 2, p. 924–935, Feb. 2019.

[68] F. Luo, T. Zhou, J. Liu, T. Guo, X. Gong, and X. Gao, "Dcenet: Diff-feature contrast enhancement network for semi-supervised hyperspectral change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 62, p. 1–14, Mar. 2024.

[69] Q. McNemar, "Note on the sampling error of the difference between correlated proportions or percentages," *Psychometrika*, vol. 12, no. 2, p. 153–157, Jun. 1947.

[70] Z. He, H. Liu, Y. Wang, and J. Hu, "Generative adversarial networks-based semi-supervised learning for hyperspectral image classification," *Remote Sens.*, vol. 9, no. 10, p. 1042, Oct. 2017.

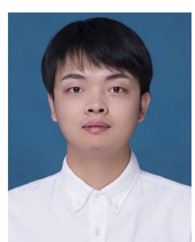
[71] E. Zhang, H. Zong, X. Li, M. Feng, and J. Ren, "Icsf: Integrating inter-modal and cross-modal learning framework for self-supervised heterogeneous change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 63, p. 1–16, 2025.

Chengle Zhou (Graduate Student Member, IEEE) received the B.S. degree in electrical engineering and automation and the M.Sc. degree in information and communication engineering from the Hunan Institute of Science and Technology, Yueyang, China, in 2019 and 2022, respectively. He is currently working toward the Ph.D. degree in geographic information science with the School of Geography and Planning, Sun Yat-Sen University (SYSU), Guangzhou, China.

Dr. Zhou was a recipient of the Outstanding Contribution Award from the Wiley Open Science Excellent Author Program in 2023 and the National Scholarship for Postgraduate granted by the Chinese Ministry of Education in 2020 and 2021.His research interests include deep learning and hyperspectral/high-resolution remote sensing image analysis.

Zhi He (Senior Member, IEEE) received the B.S. degree in information and computing science from the School of Mathematics, Harbin Institute of Technology (HIT), Harbin, China, in 2009, and the M.S. and Ph.D. degrees in control science and engineering in control science and engineering from HIT in 2011 and 2015, respectively.

In 2015, she joined the School of Geography and Planning, Sun Yat-sen University (SYSU), Guangzhou, China, where she is currently a Full Professor. Her research interests include media/street view, artificial intelligence/deep learning, and their application to wetland monitoring.



Jian Dong received the B.S. degree in information and computing sciences from East China Jiaotong University, Nanchang, China, in 2018, and the M.E. degree in computer technology with the School of Information Engineering, Jiangxi University of Science and Technology, Ganzhou, China, in 2024. He is currently pursuing the Ph.D. degree in Sun Yat-sen University, Guangzhou, China.

His research interests include wetland classification, wetland change detection, and deep learning.

Yunfei Li (Member, IEEE) received the B.S. degree from Jilin University, Changchun, China in 2018, and the M.S. degree from Sun Yat-sen University, Guangzhou, China in 2021. He is currently pursuing the Ph.D. degree in geographic information science with the School of Geography and Planning, Sun Yat-sen University (SYSU), Guangzhou, China.

His research interest is the fusion of multisource remote sensing data.

Jinchang Ren (Senior Member, IEEE) received the B.Eng., M.Eng., and D.Eng. degrees from Northwestern Polytechnical University, Xi'an, China, in 1992, 1997, and 2000, respectively, and the Ph.D. degree from the University of Bradford, Bradford, U.K., in 2019.

He is currently a Distinguished Professor with the School of Computer Sciences, Guangdong Polytechnic Normal University, Guangzhou, China. He has authored or co-authored over 400 peer-reviewed journal and conference articles. His research interests include hyperspectral image analysis, computer vision, machine learning, remote sensing and big data analytics.

Dr. Ren acts as an Associate Editor for several international journals, including IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING (TGRS) et al.

Antonio Plaza (Fellow, IEEE) received the M.Sc. degree and the Ph.D. degree in computer engineering from the Hyperspectral Computing Laboratory, the Department of Technology of Computers and Communications, University of Extremadura, Cáceres, Spain, in 1999 and 2002, respectively.

He is currently the Head of the Hyperspectral Computing Laboratory, Department of Technology of Computers and Communications, University of Extremadura. He has served as the Editor-in-Chief of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING from 2013 to 2017 and guest-edited ten special issues on hyperspectral remote sensing for different journals. His research interests include hyperspectral data processing and parallel computing of remote sensing data. Additional information: https://sites.google.com/view/antonioplaza