

AEKAN: Exploring Superpixel-Based AutoEncoder Kolmogorov-Arnold Network for Unsupervised Multimodal Change Detection

Tongfei Liu^{ID}, Member, IEEE, Jianjian Xu^{ID}, Tao Lei^{ID}, Senior Member, IEEE, Yingbo Wang^{ID}, Member, IEEE, Xiaogang Du^{ID}, Weichuan Zhang^{ID}, Member, IEEE, Zhiyong Lv^{ID}, Senior Member, IEEE, and Maoguo Gong^{ID}, Fellow, IEEE

Abstract—Multimodal change detection (MCD) has garnered significant interest due to its capacity to address a variety of emergencies in a timely and effective manner. However, discrepancies in sensors and imaging techniques often hinder the direct comparison of heterogeneous remote sensing images (HRSIs), making it difficult to extract change information. To overcome this challenge, we propose a novel superpixel-based AutoEncoder Kolmogorov-Arnold Network (AEKAN) for unsupervised MCD. The primary objective of AEKAN is to excavate the latent commonality features between HRSIs. Notably, commonality features in unchanged regions are generally more pronounced than those in changed regions, which can be leveraged to assess change magnitude. To achieve this, the proposed method utilizes the Kolmogorov-Arnold Network (KAN), renowned for its capability to model data distributions, to extract these commonality features between HRSIs. Concretely, the proposed AEKAN consists of a Siamese KAN encoder and dual KAN decoders. The Siamese encoder aims to map HRSIs and extract latent commonality features, while the dual decoders reconstruct original bitemporal images from these features. In addition, we incorporate a hierarchical commonality loss function within the Siamese encoder to train AEKAN. This loss function is designed to intentionally guide the network in capturing commonality features by minimizing the discrepancies in features extracted from HRSIs at each layer of the Siamese encoder. The extracted commonality features are then adopted to quantify the change magnitude between images through mean square error (MSE). Extensive experiments on five MCD datasets demonstrate that the proposed AEKAN

Received 11 September 2024; revised 30 October 2024; accepted 25 November 2024. Date of publication 11 December 2024; date of current version 24 December 2024. This work was supported in part by the National Natural Science Foundation of China under Grant 62271296 and Grant 62201334; in part by the Scientific Research Program Funded by Shaanxi Provincial Education Department under Grant 23JP022 and Grant 23JP014; in part by the Key Research and Development Program of Shaanxi under Grant 2024GX-YBXM-121; and in part by the Natural Science Foundation of Shaanxi, China, under Grant 2024JC-YBQN-0037. (Corresponding author: Tao Lei.)

Tongfei Liu, Jianjian Xu, Tao Lei, Yingbo Wang, Xiaogang Du, and Weichuan Zhang are with Shaanxi Joint Laboratory of Artificial Intelligence and the School of Electronic Information and Artificial Intelligence, Shaanxi University of Science and Technology, Xi'an 710021, China (e-mail: liutongfei_home@hotmail.com; 231612143@sust.edu.cn; leitao@sust.edu.cn; wangyingbo@sust.edu.cn; duxiaogang@sust.edu.cn; zwc2003@163.com).

Zhiyong Lv is with the School of Computer Science and Engineering, Xi'an University of Technology, Xi'an 710048, China (e-mail: lvzhiyong_fly@hotmail.com).

Maoguo Gong is with the Academy of Artificial Intelligence, College of Mathematics Science, Inner Mongoli Normal University, Hohhot 010028, China, and also with the Key Laboratory of Collaborative Intelligence Systems, Ministry of Education, School of Electronic Engineering, Xidian University, Xi'an 710126, China (e-mail: gong@ieee.org).

Digital Object Identifier 10.1109/TGRS.2024.3515258

outperforms existing methods. The source code is available at: <https://github.com/TongfeiLiu/AEKAN-for-MCD>.

Index Terms—Commonality features, heterogeneous images, Kolmogorov-Arnold Network (KAN), multimodal change detection (MCD).

I. INTRODUCTION

MULTIMODAL change detection (MCD) has attracted widespread attention from many scholars recently. MCD is a more challenging task in remote sensing change detection (CD) [1], [2]. The difficulty lies in using heterogeneous remote sensing images (HRSIs) that cannot be directly compared with identify land cover changes. MCD benefits from its ability to obtain land cover change information across the limitations of image modalities. For this reason, MCD can provide timely and effective data support for many practical emergency applications and land cover analysis, including landslides, floods, and earthquakes [3], [4].

In homogeneous CD, homogeneous bitemporal RSIs are usually collected from the same sensor, which has the same physical properties and imaging principles. Hence, the difference images (DIs) can be generated by measuring directly the similarity or distance between the bitemporal remote sensing images (RSIs) [2]. However, these bitemporal homogeneous RSIs are not always directly available, such as in some emergency disaster scenarios. Unlike homogeneous CD, MCD generally utilizes bitemporal HRSIs from different sensors or imaging mechanisms to detect changes, such as optical and synthetic aperture radar (SAR) images. Due to different sensors or imaging mechanisms, these HRSIs exhibit significantly different characteristics, including spectral, spatial, texture, and data distribution, which makes the unchanged areas show completely different appearances [3]. Therefore, it also leads to MCD being unable to acquire DI by directly comparing HRSIs.

To realize MCD, the intuitive idea is to extract the essential features of HRSIs, such as the real ground target attributes or sensor physical properties. Then, DI can be calculated by comparing the less affected features by the modality or imaging. Sun et al. [5], [6] proposed some methods to model the relationship between HRSIs to achieve MCD, including patch similarity graph matrix and nonlocal patch similarity. The principle of these methods is to assume that the unchanged

regions should have the same spatial structure or connection relationship. Similar methods include adaptive local structure consistency [7] and improved nonlocal patch-based graph (INLPG) structure consistency [8]. Therefore, these methods can obtain change magnitude between HRSIs by measuring the difference between such spatial structures or spatial connections. Other relevant approaches are available in [9], [10], and [11]. However, these methods inevitably present more false detections or missed pixels due to the insufficient ability of hand-crafted features to adapt to complex scenes.

In addition, another idea is to transform HRSIs from one modality to another to achieve pseudohomogeneity of bitemporal images, that is, to have similar spectra, color, appearance, so that HRSIs become comparable. According to different implementation technologies, these methods can be roughly divided into two categories: one is image regression based on the conventional image analysis, and the other is the image transformation or translation based on deep learning (DL). The conventional image analysis-based image regression can utilize the inherent structural consistency of images to regress one modality to another modality to achieve modality alignment, such as homogeneous pixel transformation [12], affinity-based image regression [13], sparse-constrained adaptive structure consistency (SCASC) [14], and image regression with structure cycle consistency (AGSCC) [15]. DL-based image transformation/translation methods convert one modality image to another modality through various convolutional neural networks (CNNs), such as convolutional autoencoders (CAEs) or generative adversarial networks (GANs), to obtain pseudohomogeneous images (X-Net [16], adversarial cyclic encoder network (ACE-Net) [16], and GAN-based approaches [17], [18], [19]). Although these recent MCD methods have made great progress, the limitation of these approaches is that the image translation quality often greatly affects the performance of MCD.

Different from the above methods, Wu et al. proposed a novel commonality autoencoder CD (CACD) approach for MCD [20]. In CACD, HRSIs are first mapped for consistency features using CAE, and then, the proposed commonality autoencoder (COAE) based on fully connected is used to extract common features. Finally, the changes between HRSIs can be detected by comparing common features. The design principle of the CACD is to assume that unchanged areas have more common features than changed areas, and these features can be employed to identify changes by extracting and comparing common features. The CACD has proven this assumption and achieved effective MCD. Nevertheless, three aspects of CACD limit the accuracy of MCD. First, since CACD is a sliding window-based method, different window sizes may introduce other objects and interfere with common feature extraction, thus causing false detection or missed detection. Second, CACD is a two-stage training framework, which makes the training process more complicated. Third, although CAE maps HRSIs to more consistent feature representations to weaken the impact of modal differences, it also leads to a lot of information loss. At the same time, it may also cause the loss of some effective common features in COAEs directly. Furthermore, the ability of COAEs based on fully connected to represent common features may be difficult to adapt to

complex HRSIs due to limited model capacity and capabilities. Therefore, overcoming these limitations will help to improve the performance of MCD further.

Inspired by CACD [20], the motivation of our approach lies in the following two points. On the one hand, the defect of the sliding window-based method that is easily disturbed by other objects can be avoided by using superpixels as analysis units. Meanwhile, suppose the network is designed as a one-stage training process to extract commonality features. In that case, it can alleviate the loss of information between HRSIs and help acquire more effective and rich commonality features. On the other hand, since the capacity and capabilities of a simple fully connected network are limited, introducing a network with stronger modeling capabilities can further enhance the ability to capture the commonality features.

According to the above motivations, we proposed a superpixel-based AutoEncoder Kolmogorov-Arnold Network (AEKAN) for MCD in HRSIs. The major contributions of this article are summarized as follows.

- 1) To the best of the authoers' knowledge, this is the first time that a pure KAN model has been proposed for MCD. The effectiveness and capabilities of the pure KAN model for MCD have been verified, providing a reference for future research focused on reliable and interpretable MCD methods.
- 2) We proposed a superpixel-based AEKAN for unsupervised MCD, which can capture the latent commonality features between HRSIs. This study has verified and visualized these commonality features, showing their utility in directly acquiring high-quality DIs and addressing the challenges of feature incomparability between HRSIs in MCD.
- 3) A hierarchical commonality loss is introduced within the Siamese KAN encoder, which can facilitate the commonality feature extraction between HRSIs by minimizing the discrepancies in features extracted from HRSIs at each layer of the Siamese encoder.
- 4) Comprehensive comparative and ablation experiments conducted on five challenging MCD datasets affirm the effectiveness and superiority of the proposed AEKAN.

The rest of this article is arranged as follows. Related works are reviewed in Section II. Section III describes the proposed AEKAN in detail. In Sections IV and V, experimental results and discussion are provided, respectively. Finally, the conclusion and future works are summarized in Section VI.

II. RELATED WORKS

A. MCD Methods

In the recent decade, researchers have proposed many solutions for MCD. These methods can be roughly classified into two categories: traditional image analysis-based and DL-based.

1) *Traditional Image Analysis-Based MCD Methods:* In the early period, some methods exploit hand-crafted modality-independent features to calculate DI directly. Mercier et al. [21] believe that there is a potential relationship between HRSIs and propose using the copula theory to model them.

Then, the differences between HRSIs are obtained by comparing the statistics based on Kullback–Leibler. Prendes et al. [22] proposed a similarity measurement method based on the sensor physical properties for MCD. The approach adopts manifold learning to capture the noise model and local joint distribution of HRSIs and is used to detect differences between HRSIs. Other early MCD methods include [23] and [24]. In recent years, some methods based on the machine learning and mathematical models have been developed for MCD. Touati et al. [25] applied a multidimensional scaling representation strategy to map HRSIs into the same feature space, which enables the use of any homogeneous CD methods to compare the statistical features between HRSIs to obtain changes. Mignotte [26] exploit the fractal projection based on spatial self-similarity encoding to project images from one modality into another for MCD. Also, the dictionary learning can detect changes by building coupled dictionaries to establish relationships between HRSIs [27], [28]. In addition, the graph signal processing has been applied in MCD. A graph-based fusion CD approach (GBF-CD) [29] and a graph learning based on signal smoothness (GLSS) [30] are proposed in homogeneous and heterogeneous RSIs. Recently, Sun et al. [31], [32] further proposed a series of MCD methods based on the graph structure analysis. For instance, an iterative robust graph and Markovian co-segmentation (IRG-McS) approach [31] and an iterative structure transformation and conditional random field (IST-CRF) approach [32] are presented, which can obtain changes by comparing graph structure with modality-invariant. Similarly, a method named graph based image regression and Markov random field segmentation method (GIR-MRF) was devised in [33]. This approach regresses one modality image to another modality through graph structure consistency to make HRSIs comparable and then adopts the Markov segmentation model to obtain the change image. Other advanced image analysis methods can also be found in [34] and [35].

2) *DL-Based MCD Methods*: With the development of DL technologies, more novel approaches have been developed for MCD in HRSIs [1], [3], [36]. These DL-based methods have provided more perspectives for achieving MCD in recent years.

The first solution is that DL methods utilize powerful learning capabilities to achieve MCD by selecting labeled sample supervision [37], [38]. However, such methods require manual labeling of training samples, which is greatly limited for small-sized HRSIs. To alleviate the dependence on labeled samples, some methods present a variety of strategies to select pixels that can represent changes and unchanged pixels from HRSIs as training samples to realize unsupervised MCD, such as sample selection strategies based on self-paced learning [39]. Likewise, Shi et al. [40] proposed self-guided autoencoders to choose reliable pseudolabels and obtained more reliable pseudo labels for training classifiers through ongoing iterations.

The second solution is to employ the powerful image generation and transformation capabilities of DL to transform HRSIs into pseudohomogeneous images for achieving MCD forcibly, such as affinity-based change prior [16], Cycle GANs [41], bipartite graph attention autoencoders based on an image translation framework [42], and dynamic multilevel

feature fusion network [43]. Zhan et al. [44] proposed an MCD method for optical and SAR images. This method approximates the data statistical distribution of optical images by logarithmic transformation of SAR images and trains the network through joint feature learning to realize MCD. In [45], a bipartite adversarial autoencoder network with clustering (BAACL) is devised for MCD. The BAACL can perform style alignment on the HRSIs through the structural consistency loss and adversarial loss constraints to achieve MCD.

The third approach is to map HRSIs into a feature space to enhance the comparability between HRSIs to accomplish MCD. For example, Liu et al. [46] first devised a symmetric convolutional coupling network (SCCN) for MCD, which can obtain more consistent feature representation by mapping HRSIs to the same feature space, thereby effectively implementing MCD. A conditional adversarial network (CANet) is designed for MCD in [47], which aims to map HRSIs to a certain feature space using conditional GANs to make the features of HRSIs comparable. This type of approach can also be found in [48] and [49]. In addition, the fourth option is to capture modality-independent features through the representation ability of network abstract features to measure the changes between HRSIs, such as CACD [20]. Also, Chen et al. [50], [51] promoted a structural relationship graph CAE (SR-GCAE) method and a Fourier domain structural relationship analysis method for unsupervised MCD, which has the capability to acquire DI by calculating the similarity of structural relationships [50], [51].

B. Kolmogorov-Arnold Networks

Nowadays, in addition to the popular CNN and Transformer, various new networks are emerging to explore new structures and models with more interpretability and high performance, such as Kolmogorov-Arnold networks (KANs) [52]. The mathematical principle of KAN is the Kolmogorov-Arnold representation theorem; that is, multivariate variables can be represented by the superposition of univariate continuous functions with two parameters. In particular, KAN's activation function is learnable, providing greater data-fitting possibilities. Therefore, the major competitiveness of KAN [52] lies in its greater interpretability and a stronger ability to fit data distribution with only a few parameters compared with traditional multilayer perception (MLP). In recent months, KAN has begun to attract attention in the field of remote sensing, such as satellite image scene classification [53], hyperspectral image classification [54], [55], and hyperspectral image CD [56]. Concretely, Cheon [53] proposed a KAN for satellite scene image classification, which integrates KAN and CNN by adopting KAN instead of MLP to improve the classification accuracy. A hybrid KAN-based model and a HyperKAN are devised for hyperspectral image classification in [54] and [55]. These methods demonstrate the ability of KAN to extract features from complex hyperspectral data effectively. It is exciting that a SpectralKAN has recently been proposed for hyperspectral CD [56]. In the SpectralKAN [56], a spectral KAN and a spatial KAN are applied to capture spectral and spatial features for identifying changes from hyperspectral images. The preliminary exploration of these methods has

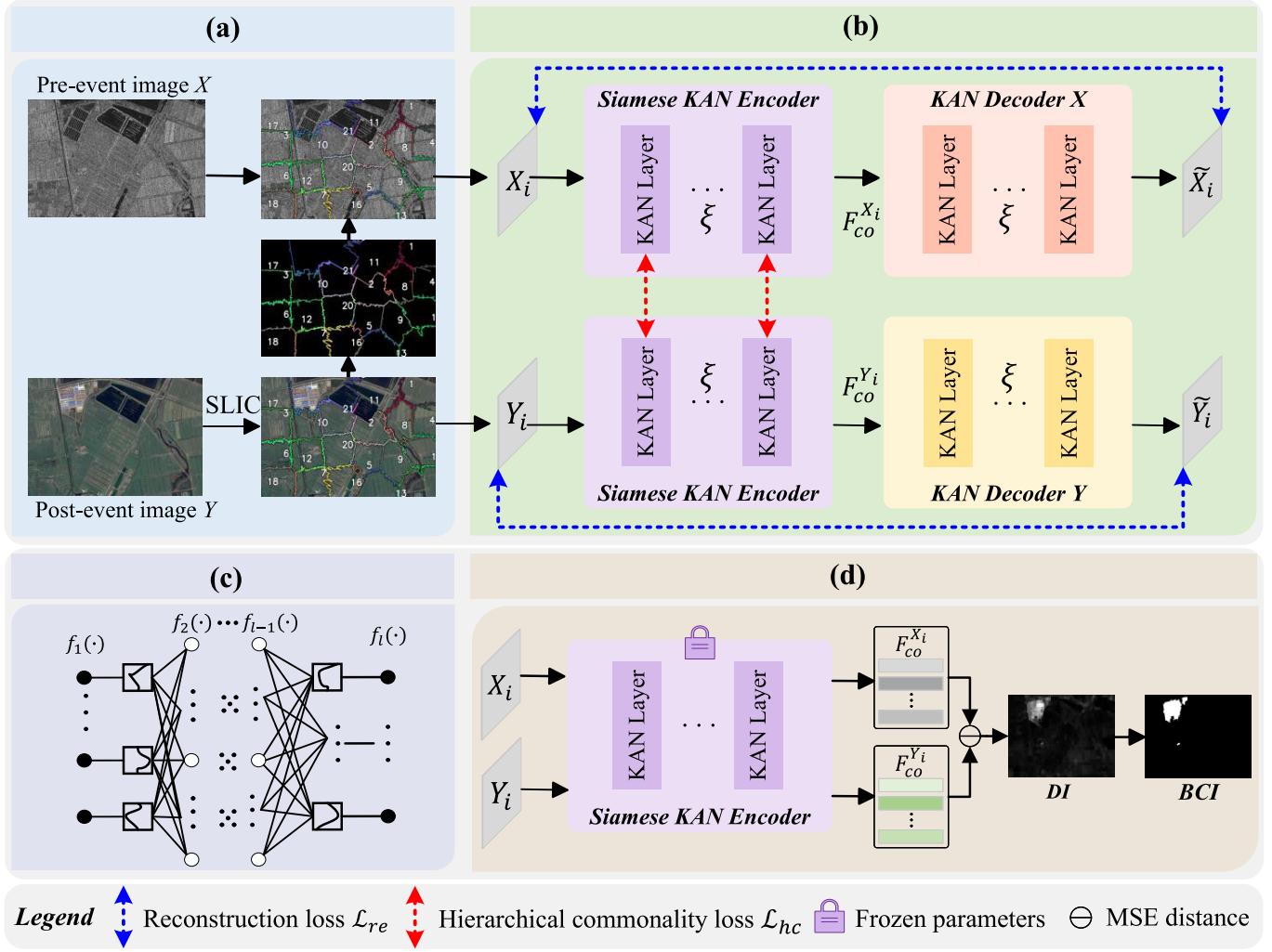


Fig. 1. Framework of the proposed superpixel-based AEKAN for unsupervised MCD. (a) Data preprocessing: obtaining superpixels for training and testing for the proposed AEKAN by using the SLIC algorithm. (b) AEKAN: capturing the commonality features between HRSIs by minimizing the reconstruction loss \mathcal{L}_{re} and the hierarchical commonality loss \mathcal{L}_{hc} . In addition, X_i and Y_i represent the original superpixels of HRSIs, respectively, and \tilde{X}_i and \tilde{Y}_i represent the reconstructed superpixels of HRSIs, respectively. $F_{co}^{X_i}$ and $F_{co}^{Y_i}$ represent the latent commonality features of HRSIs output by the Siamese encoder, respectively. (c) Generating DI and BCI: measuring change magnitude by MSE distance between the commonality features of HRSIs, and using Otsu [57] to divide DI into BCI. (d) Illustration of the KAN layer. $f_l(\cdot)$ represents the activation function of the l th layer.

presented that KAN has the potential for a variety of remote sensing tasks. Driven by this, this article aims to continue exploring the effectiveness and superiority of KAN on MCD.

III. METHODOLOGY

A. Overview of the Proposed AEKAN

Since MCD cannot directly compare HRSIs to obtain the changes between them, this article aims to excavate the potential commonality features between HRSIs to measure the changes between them directly. To achieve this goal, this article proposes a superpixel-based AEKAN for unsupervised MCD, as shown in Fig. 1. The proposed approach includes three steps to implement MCD. First, the data preprocessing (superpixel segmentation and normalization) is required in the proposed superpixel-based AEKAN. That is, the simple linear iterative clustering (SLIC) algorithm is used to obtain superpixels from the postevent image, and the bitemporal HRSIs are normalized by the min–max normalization algorithm, respectively. Second, superpixels are used as processing units in the proposed framework, and they are input into the AEKAN composed of a Siamese KAN encoder and dual KAN decoders for

reconstruction to train AEKAN. Based on the Siamese KAN encoder, the latent commonality features between HRSIs can be gradually learned by minimizing the proposed hierarchical commonality loss. Finally, the commonality features between HRSIs can be captured by inference of the trained Siamese KAN encoder. Then, the DI can be calculated by measuring directly the distance between the commonality features of bitemporal HRSIs, and the DI is divided into the binary change image (BCI) through the Otsu [57] algorithm. According to this, the proposed superpixel-based AEKAN will be described in detail as follows.

B. Data Preprocessing

In this work, the proposed superpixel-based AEKAN requires two data preprocessing operations, i.e., normalization and superpixel segmentation. Details of data preprocessing are as follows.

First, the image normalization is required in the proposed superpixel-based AEKAN because it can transform HRSIs into a more consistent data representation, eliminating the adverse effects of different dimensions between HRSIs. Specifically,

we directly use min–max normalization to preprocess optical images. For SAR images, a logarithmic transform is first required to suppress speckle noise, and then, the transformed image is also normalized using min–max normalization. Here, the normalized preevent image and postevent image can be defined as $X \in \mathbb{R}^{H \times W \times C}$ and $Y \in \mathbb{Z}^{H \times W \times C}$, where H , W , and C denote the height, width, and channel of HRSIs, respectively. Also, \mathbb{R} and \mathbb{Z} refer to the modalities of the pre- and post-event images, respectively. Except for image segmentation, all other processes use normalized X and Y in our proposed method.

Second, to alleviate the limitations of sliding windows, we use superpixels as analysis units to suppress the individual features of different objects and improve the ability to represent the latent commonality features between HRSIs. The process of superpixel segmentation is shown in Fig. 1(a). The superpixel segmentation is performed on the original postevent image by the SLIC algorithm from the skimage package to obtain superpixel index $\Omega = \{\Omega_1, \Omega_2, \dots, \Omega_T\}$, where T is the actual number of superpixels. Then, based on the segmentation index Ω , we can acquire the superpixel lists of X and Y as $X_\Omega \in \mathbb{R}^{H \times W \times C} = \{X_i | 1 \leq i \leq T\}$ and $Y_\Omega \in \mathbb{Z}^{H \times W \times C} = \{Y_i | 1 \leq i \leq T\}$, respectively. Hence, X_Ω and Y_Ω are utilized as the input units of the proposed AEKAN for training and testing.

Notably, our method only involves two parameters in the SLIC superpixel segmentation algorithm that need to be set to adapt to datasets of different sizes and object coverage. These two parameters are $n_segments$ and compactness in the SLIC algorithm, which denotes the number of superpixel and segmentation compactness, respectively. The segmentation granularity of superpixels may affect the performance of the proposed superpixel-based AEKAN because different MCD datasets often cover different types and scales of objects. Therefore, in the proposed approach, the parameters $n_segments$ (\mathcal{P}_n) and compactness (\mathcal{P}_c) need to be adjusted to adapt to different MCD datasets. Subsequent experiments will analyze the parameter sensitivity on different MCD datasets.

C. Proposed AEKAN

In this work, we introduce the latest KAN inspired by the Kolmogorov-Arnold representation theorem to improve the commonality feature representation ability in the proposed AEKAN. Herein, we first briefly illustrate the principle of the KAN layer and then present the structure and process of the proposed AEKAN in detail.

1) Description of KAN Layer: The KAN is a type of neural network architecture based on the Kolmogorov-Arnold representation theorem. This theorem states that any multivariable continuous function can be a combination of a finite number of univariate continuous functions. The KAN can overcome the limitations of this theorem by introducing learnable activation functions, allowing the network to have any number of layers and nodes. As shown in Fig. 1(d), in the KAN, each layer's activation function is represented by a function vector, with each neuron having a different activation function. Herein, we can assume that the input to the l th layer of KANs is

$\mathbf{z}_l \in \mathbb{G}^{m_l}$, and the output is $\mathbf{z}_{l+1} \in \mathbb{G}^{m_{l+1}}$, \mathbf{z}_{l+1} can be computed as follows:

$$\mathbf{z}_{l+1} = f_l(\mathbf{z}_l) \quad (1)$$

where f_l is the activation function matrix of the l th layer, that is,

$$f_l(\cdot) = \begin{pmatrix} f_{l,1,1}(\cdot) & f_{l,1,2}(\cdot) & \cdots & f_{l,1,m_l}(\cdot) \\ f_{l,2,1}(\cdot) & f_{l,2,2}(\cdot) & \cdots & f_{l,2,m_l}(\cdot) \\ \vdots & \vdots & \ddots & \vdots \\ f_{l,m_{l+1},1}(\cdot) & f_{l,m_{l+1},2}(\cdot) & \cdots & f_{l,m_{l+1},m_l}(\cdot) \end{pmatrix}. \quad (2)$$

The activation function $f(\cdot)$ in KANs consists of two parts: a weighted activation function and a spline function, and it can be defined as follows:

$$f(y) = \varphi_a \beta(y) + \varphi_b \mathcal{S}(y) \quad (3)$$

where $\beta(\cdot)$ typically refers to the sigmoid linear unit (SiLU) function. The spline function $\mathcal{S}(y)$ is composed of multiple B -spline basis functions, that is,

$$\mathcal{S}(y) = \sum_{j=0}^t d_j B_j^r(y) \quad (4)$$

where $B_j^r(y)$ is the j th B -spline basis function, d_j is the corresponding weight, t is the number of knot spans determined by the knot vector, and r is the degree of the spline function. By appropriately choosing the knot vector, degree r , and weights d , the spline function can approximate any continuous function. φ_a and φ_b are the learnable weights of the $\beta(\cdot)$ and spline functions, respectively. Based on this, the approach allows KANs to learn and represent complex functional relationships flexibly. As a result, in the proposed approach, constructing a KAN-based autoencoder is expected to extract effectively the potential commonality features between HRSIs.

2) Structure of AEKAN: Based on the above KAN layer, we designed the proposed AEKAN, and its structure is illustrated in Fig. 1(b). The proposed AEKAN contains three parts, i.e., a Siamese KAN encoder and dual KAN decoders, both of which consist of ξ pure KAN layers. Here, we can define the Siamese KAN encoder as $\mathcal{KE}(\cdot)$ and dual KAN decoders as $\mathcal{KD}_X(\cdot)$ and $\mathcal{KD}_Y(\cdot)$, respectively. According to this, the process of the proposed AEKAN is described as follows. First, each pair of superpixels X_i and Y_i in X_Ω and Y_Ω are input into the Siamese KAN encoder $\mathcal{KE}(\cdot)$ to capture more consistent commonality features. They can be described as follows:

$$F_{co}^{X_i} = \mathcal{KE}(X_i), \quad F_{co}^{Y_i} = \mathcal{KE}(Y_i) \quad (5)$$

where $F_{co}^{X_i}$ and $F_{co}^{Y_i}$ are the latent commonality features of X_i and Y_i , respectively. Therefore, the Siamese KAN encoder $\mathcal{KE}(\cdot)$ aims to map HRSIs into the same feature space for more effective commonality feature representation. Then, based on $F_{co}^{X_i}$ and $F_{co}^{Y_i}$, X_i and Y_i can be reconstructed through the dual KAN decoders $\mathcal{KD}_X(\cdot)$ and $\mathcal{KD}_Y(\cdot)$, that is,

$$\widetilde{X}_i = \mathcal{KD}_X(F_{co}^{X_i}), \quad \widetilde{Y}_i = \mathcal{KD}_Y(F_{co}^{Y_i}) \quad (6)$$

where \widetilde{X}_i and \widetilde{Y}_i refer to the pseudoreconstructed X_i and the pseudoreconstructed Y_i , respectively. Hence, our designed

AEKAN is able to excavate the latent commonality features between HRSIs in an unsupervised learning manner through reconstruction tasks.

In this situation, we propose a series of loss functions to train the proposed AEKAN in an unsupervised learning manner. Concretely, the proposed loss function includes two subitems: reconstruction loss \mathcal{L}_{re} and hierarchical commonality loss \mathcal{L}_{hc} . The reconstruction loss \mathcal{L}_{re} aims to reconstruct the original superpixels X_i and Y_i into \tilde{X}_i and \tilde{Y}_i , respectively, and satisfy: $\tilde{X}_i = \mathcal{KD}_X(\mathcal{KE}(X_i)) \approx X_i$ and $\tilde{Y}_i = \mathcal{KD}_Y(\mathcal{KE}(Y_i)) \approx Y_i$. Consequently, the reconstruction loss \mathcal{L}_{re} can be marked as follows:

$$\mathcal{L}_{\text{re}} = \|X_i - \tilde{X}_i\|_{\text{MSE}} + \|Y_i - \tilde{Y}_i\|_{\text{MSE}} \quad (7)$$

where $\|\cdot\|_{\text{MSE}}$ represents a mean square error (MSE), which can measure the similarity between the original superpixel and pseudoreconstructed superpixel. Therefore, X_i and Y_i can be approximately reconstructed as \tilde{X}_i and \tilde{Y}_i by minimizing the reconstruction loss \mathcal{L}_{re} .

However, it is a difficult task to effectively guide the proposed AEKAN to extract the commonality features between HRSIs using only the reconstruction loss \mathcal{L}_{re} . To this end, we further propose a hierarchical commonality loss \mathcal{L}_{hc} , which aims to explicitly guide the Siamese KAN encoder in the proposed AEKAN to obtain the latent commonality features between HRSIs. Herein, the hierarchical commonality loss \mathcal{L}_{hc} can be calculated as follows:

$$\mathcal{L}_{\text{hc}} = \sum_{k=1}^{\xi} \|F_{\text{co}}^{X_i}(k) - F_{\text{co}}^{Y_i}(k)\|_{\text{MSE}} \quad (8)$$

where ξ is the number of KAN layers in the Siamese KAN encoder and $F_{\text{co}}^{X_i}(k)$ and $F_{\text{co}}^{Y_i}(k)$ refer to the commonality features of X_i and Y_i at the k th layer, respectively. Accordingly, the effective commonality features between HRSIs can be enhanced by minimizing the MSE distance between $F_{\text{co}}^{X_i}(k)$ and $F_{\text{co}}^{Y_i}(k)$. Hence, our proposed total loss can be represented as follows:

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{re}} + \mathcal{L}_{\text{hc}} \quad (9)$$

where $\mathcal{L}_{\text{total}}$ refers to the total loss in the proposed approach. To this end, we can extract and enhance the commonality features between HRSIs in an unsupervised manner by minimizing the total loss $\mathcal{L}_{\text{total}}$. Finally, the change magnitude between HRSIs can be measured by directly computing the distance between the commonality features of HRSIs.

D. Generating DI and BCI

When the proposed AEKAN is trained, a trained Siamese KAN encoder $\mathcal{KE}(\cdot)$ and dual KAN decoders $\mathcal{KD}_X(\cdot)$ and $\mathcal{KD}_Y(\cdot)$ can be obtained. Also, we can use the trained AEKAN to extract the commonality features between HRSIs. Based on this, DI and BCI can be generated at this stage. Specifically, in the proposed AEKAN, we adopt the commonality features $F_{\text{co}}^{X_i}(\xi)$ and $F_{\text{co}}^{Y_i}(\xi)$ output by the last layer of the Siamese KAN encoder to measure the change magnitude between each pair of superpixels within HRSIs. Based on this, DI between

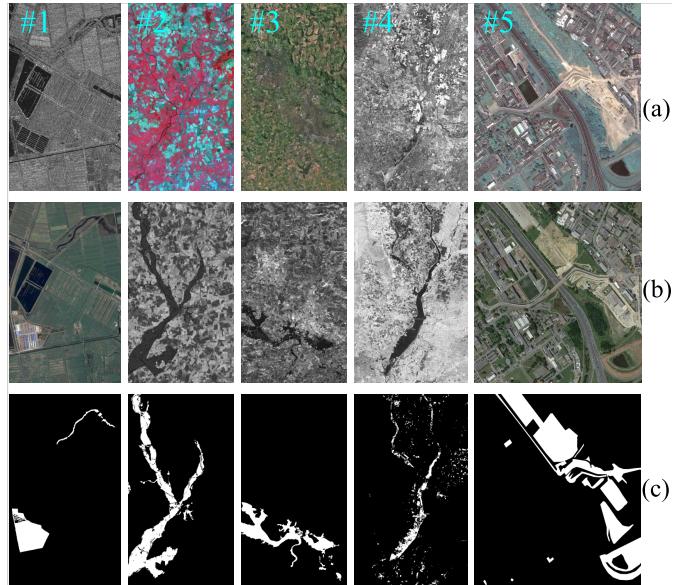


Fig. 2. Datasets #1–#5. (a) Pre-event images. (b) Postevent images. (c) Ground reference images.

HRSIs can be generated until all superpixels are processed in this manner. This can be expressed as follows:

$$\mathbb{D} = \{\mathbb{D}_i = \|F_{\text{co}}^{X_i}(\xi) - F_{\text{co}}^{Y_i}(\xi)\|_{\text{MSE}}, i \in \{1 \leq i \leq T\}\} \quad (10)$$

where \mathbb{D}_i represents the difference between each superpixel pair within HRSIs and \mathbb{D} represents the final DI of the proposed superpixel-based AEKAN approach. For this, a promising \mathbb{D} can be acquired. Accordingly, the latent commonality features between HRSIs can be effectively captured in the proposed superpixel-based AEKAN, and the differences between HRSIs can be measured by calculating the distance between their commonality features. Finally, a simple and widely used unsupervised threshold segmentation algorithm, Otsu [57], is employed to divide the DI into a BCI.

IV. EXPERIMENTS AND RESULTS

In this section, we will describe the experimental settings in detail, including the datasets and evaluation metrics. Then, comparative methods and experimental setups are introduced. Afterward, we will present the experimental results and analysis for DI and BCI to validate the effectiveness and superiority of the proposed AEKAN. The details are as follows.

A. Datasets Descriptions and Evaluation Metrics

1) *Datasets Descriptions*: In the experiments, five public and challenging MCD datasets are selected to verify the effectiveness and superiority of the proposed AEKAN. The detailed information of these five datasets #1–#5 are shown in Fig. 2 and Table I. The spatial resolution of these datasets ranges from 0.52 to 25 m, and they are captured in cities in different countries and involve different land cover change events, such as floods and building reconstruction. Importantly, these MCD datasets are captured by different sensors. Among them, the HRSIs of datasets #1, #3, and #4 are captured by different optical and SAR sensors, and the HRSIs of datasets #2 and #5 are obtained by different optical sensors. Therefore, these MCD datasets are difficult to compare directly to achieve

TABLE I
DETAILED DESCRIPTIONS OF THE FIVE MCD EXPERIMENTAL DATASETS IN THE EXPERIMENTS

Datasets	Date	Modality	Size (used band)	Spatial resolution (m)	Change Event	Location
#1	June 2008/Sept. 2012	Radarsat-2/Google Earth	593 × 921 × 1/3 (all)	8	Constructions	Shuguang, China
#2	1999/2000	Spot/NDVI	990 × 554 × 3/1 (all)	≈ 25	Flooding	Gloucester, England
#3	July 2006/July 2007	QuickBird-2/TerraSAR-X	4135 × 2325 × 3/1 (all)	≈ 0.65	Flooding	Gloucester, England
#4	Jan. 2017/Feb. 2017	Landsat-8/Sentinel-1A	875 × 500 × 11/3 (NIR/SAR)	≈ 15	Flooding	Sutter County, USA
#5	May 2012/July 2013	Pleiades/WorldView2	2000 × 2000 × 3/3 (all)	0.52	Constructions	Toulouse, France

MCD due to the differences in imaging principles. To this end, these five MCD datasets are adopted in our experiments to verify the superiority of the proposed AEKAN.

2) *Evaluation Metrics*: To better quantitatively evaluate the performance and superiority of the proposed AEKAN and the compared methods, we chose different evaluation criteria for DI and BCI. For DI, we selected the receiver operating characteristics (ROCs) curve and the precision-recall (PR) curve to intuitively analyze the quality of DI. Furthermore, the areas under ROC curves (AURs) and the areas under PR curves (AUPs) can be exploited to quantitatively compare the quality of DI generated by different methods. For BCIs, three popular evaluation criteria are adopted in the experiments, including overall accuracy (OA), Kappa coefficient (Ka), and *F*1-score (*F*1). These evaluation metrics have been widely used in CD [15], [20]. We will not elaborate here.

B. Benchmark Methods and Experimental Setup

1) *Benchmark Methods*: Herein, we selected many representative and advanced methods as benchmarks for comparative experiments, including image analysis-based MCD and DL-based methods. Concretely, the image analysis-based MCD approaches include IRG-McS [31], GIR-MRF [33], SCASC [14], AGSCC [15], IST-CRF [32], GBF-CD [29], and GLSS [30]. The DL-based methods include CANet [47], X-Net [16], ACE-Net [16], CACD [20], SR-GCAE [50], and BAACL [42]. These benchmark methods provide multiple perspectives for evaluating the effectiveness of CD. Based on these, we can comprehensively verify the feasibility and robustness of the proposed AEKAN by conducting comparative experiments with these methods to demonstrate its superiority in MCD.

2) *Experimental Setup*: In our experiments, the parameters of different methods need to be set. To ensure the fairness of the comparison experiments, the parameter settings of the compared methods are consistent with their original paper or the best parameters obtained through trial and error. In the proposed AEKAN, the parameters n_{segments} (\mathcal{P}_n) and compactness (\mathcal{P}_c) in the SLIC segmentation algorithm is set as dataset #1: $\mathcal{P}_n = 1100$ and $\mathcal{P}_c = 20$; dataset #2: $\mathcal{P}_n = 1700$ and $\mathcal{P}_c = 36$; dataset #3: $\mathcal{P}_n = 1600$ and $\mathcal{P}_c = 40$; dataset #4: $\mathcal{P}_n = 2000$ and $\mathcal{P}_c = 0.35$; and dataset #5: $\mathcal{P}_n = 1400$ and $\mathcal{P}_c = 38$. Notably, the parameter sensitivity analysis will be presented and discussed in Section V. In addition, in the experiments, two KAN layers are deployed in both the Siamese KAN encoder $\mathcal{KE}(\cdot)$ and the dual decoders $\mathcal{KD}_X(\cdot)$ and $\mathcal{KD}_Y(\cdot)$ in the proposed AEKAN. Based on this, the proposed superpixel-based AEKAN was executed on

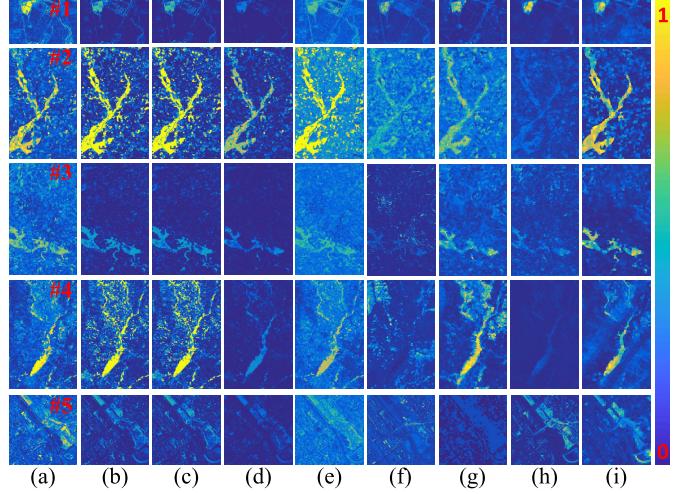


Fig. 3. Comparison of DI visualization with different methods for datasets #1–#5. (a) IRG-McS [31], (b) GIR-MRF [33], (c) SCASC [14], (d) AGSCC [15], (e) IST-CRF [32], (f) CACD [20], (g) SR-GCAE [50], (h) BAACL [42], and (i) proposed AEKAN.

a workstation configured with an Intel¹ Xeon¹ Gold 6226R with a clock speed of 2.90 GHz and an NVIDIA GeForce RTX 3090 GPU. Also, the proposed superpixel-based AEKAN was deployed on the PyTorch framework. The AdamW optimizer is adopted to minimize the loss $\mathcal{L}_{\text{total}}$, and the learning rate and weight decay are both set to 0.0001. According to these settings, we performed the experiments.

C. Comparison on DIs

In the proposed superpixel-based AEKAN, DI can be generated, which can effectively reflect the change amplitude between HRSIs in the MCD data. Therefore, a high-quality DI can ensure the accuracy and reliability of CD results. To this end, to verify the effectiveness and superiority of DI, we intuitively display and compare DI from multiple perspectives, including the heatmaps of DIs (Fig. 3), and their PR and ROC curves (Fig. 4). Here, we selected methods that can generate DI from the benchmark methods for comparison, including IRG-McS [31], GIR-MRF [33], SCASC [14], AGSCC [15], IST-CRF [32], CACD [20], SR-GCAE [50], and BAACL [42].

1) *Visual Analysis and Comparison*: To achieve a more intuitive visual analysis, we colorized the DI using a gradient of colors from 0 to 1 to represent the change magnitude between HRSIs. In addition, to better verify the quality of the DI generated by the proposed AEKAN, we compared the color DI generated by the proposed method with those by the comparison methods. The DIs of these comparative approaches on the five MCD datasets are shown in Fig. 3. By observing visual

¹Registered trademark.

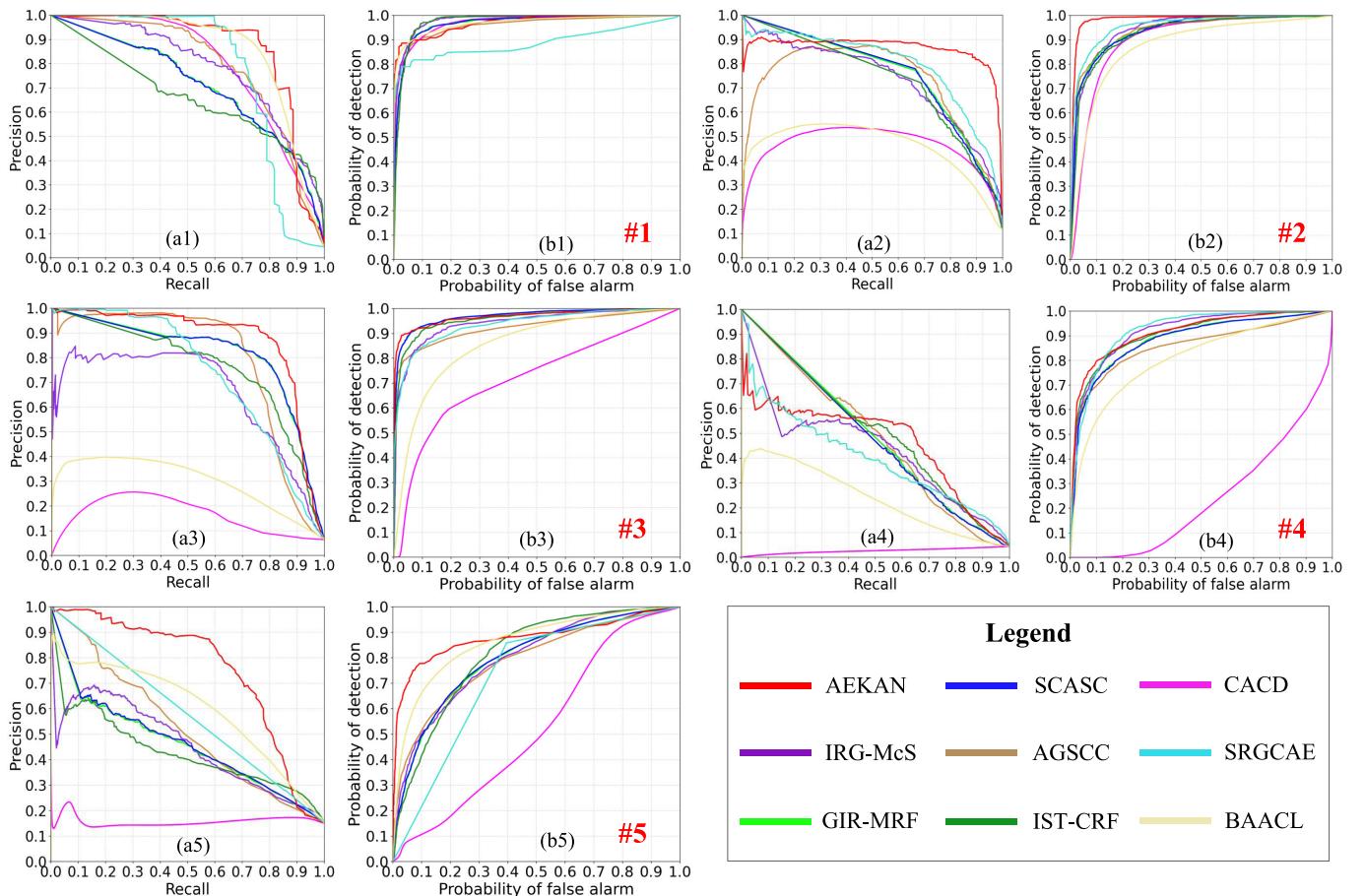


Fig. 4. Comparisons of ROC and PR curves of DI generated with different methods for datasets #1–#5. (a1)–(a5) PR curves. (b1)–(b5) ROC curves.

DIs, the proposed AEKAN exhibits a better color contrast on almost all datasets, making changed areas more prominent and easily identifiable. Furthermore, compared with other methods, our method generated results with a reasonably smooth color gradient process on datasets #2–#4, which demonstrates that our AEKAN network maintains feature continuity and accuracy during the feature extraction process. Most notably, our method generates images with significantly less noise and a relatively cleaner background. Therefore, by combining AEKAN and the hierarchical commonality loss constraint in the proposed AEKAN, our method performs well in DI quality, demonstrating its advantages in MCD.

2) *ROC and PR Curves Comparison:* To further evaluate the performance of DI, we used PR and ROC curves and the area under the curve AUP and AUC for qualitative and quantitative analyses. As presented in Fig. 4, the PR and ROC curves of different methods on datasets #1–#5 can be obtained. By comparison, it can be seen that our method performs better on both the ROC and PR curves of most datasets. Especially on datasets #2 and #5, our method is significantly close to the upper left corner in the ROC curve, while on datasets #3 and #5, the PR curve is also significantly close to the upper right corner. This shows that compared with other methods, our approach can better measure the variation range between HRSIs by calculating the distance between the commonality features of HRSIs to obtain DI, making DI have better separability. In addition, the area under the curve (AUC) and AUP are also computed to compare model performance quantitatively, as listed in Table II. As can be seen from Table II, the proposed

AEKAN acquires a better performance, with average AUC and AUP on datasets #1–#5 being 0.026 and 0.101 higher than other second-best compared methods, respectively. Hence, the quantitative comparison of AUC and AUP also supports the conclusion of the DI visual comparison.

D. Comparison on BCIs

In addition to the DI comparison, the BCIs of different methods are also compared to verify the performance and the superiority of the proposed AEKAN in this experiment. The BCIs and accuracy of the proposed AEKAN and benchmark methods can be acquired, as shown in Fig. 5 and Table III. By observing visual BCIs (see Fig. 5), some methods are more sensitive to noise and can easily misjudge noise as changing areas, which increases the false detections, such as GBF-CD [29], GLSS [30], CACD [20], and SR-GCAE [50]. Our proposed AEKAN presents fewer false detections and missed detections on almost all datasets, and the background is cleaner. This shows that the proposed AEKAN can effectively detect the changes between HRSIs by extracting the commonality features of HRSIs, and effectively reduce the false detection caused by the modal differences of HRSIs.

Table III provides the quantitative evaluation results of different methods on datasets #1–#5, and the red and blue represent the best and second-best evaluation accuracies, respectively. Notably, since X-Net [16] and ACE-Net [16] cannot achieve effective detection on dataset #3, no corresponding quantitative results are provided in this experiment. As listed in Table III, we can see that our proposed method performs

TABLE II
COMPARISONS BETWEEN THE PROPOSED AEKAN AND OTHER RELATED METHODS ON AUC AND AUP FOR DATASETS #1–#5. THE BEST RESULTS AND SECOND-BEST ACCURACY ARE PRESENTED IN RED AND BLUE, RESPECTIVELY

Methods	Dataset #1		Dataset #2		Dataset #3		Dataset #4		Dataset #5		Average	
	AUC	AUP										
IRG-McS [31]	0.980	0.785	0.946	0.735	0.941	0.659	0.917	0.433	0.799	0.454	0.917	0.613
GIR-MRF [33]	0.967	0.712	0.939	0.759	0.968	0.816	0.885	0.489	0.798	0.459	0.911	0.647
SCASC [14]	0.967	0.710	0.938	0.758	0.968	0.816	0.884	0.483	0.799	0.463	0.911	0.646
AGSCC [15]	0.961	0.791	0.939	0.716	0.926	0.794	0.859	0.484	0.791	0.524	0.895	0.662
IST-CRF [32]	0.974	0.663	0.928	0.734	0.957	0.761	0.908	0.514	0.813	0.429	0.916	0.620
CACD [20]	0.971	0.796	0.902	0.459	0.720	0.163	0.240	0.026	0.524	0.158	0.675	0.320
SR-GCAE [50]	0.888	0.777	0.957	0.795	0.937	0.738	0.921	0.402	0.731	0.578	0.888	0.658
BAACL [42]	0.970	0.855	0.875	0.458	0.856	0.300	0.807	0.238	0.863	0.597	0.874	0.490
Proposed AEKAN	0.970	0.863	0.984	0.865	0.968	0.873	0.914	0.461	0.879	0.754	0.943	0.763

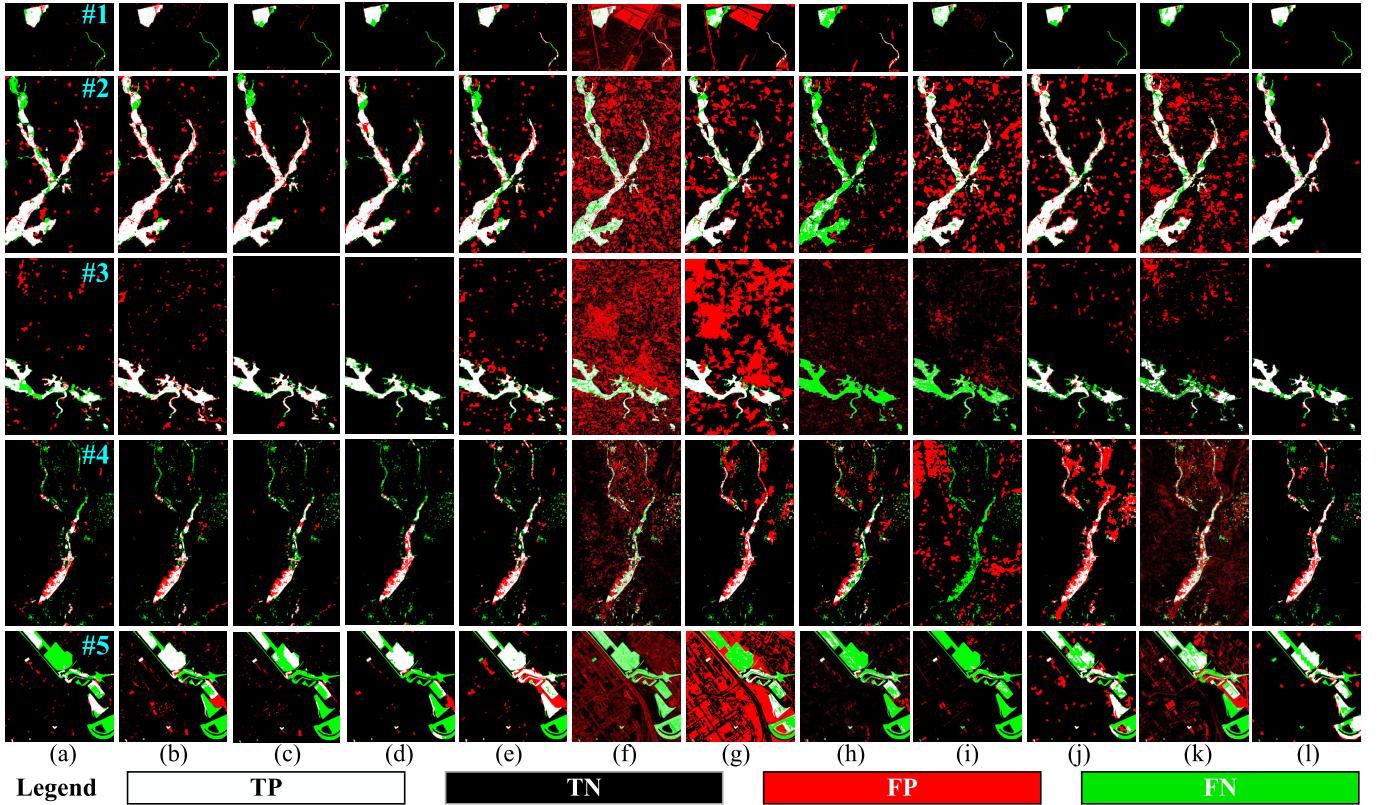


Fig. 5. BCIs of different methods on five MCD datasets #1–#5. (a) IRG-McS [31], (b) GIR-MRF [33], (c) SCASC [14], (d) AGSCC [15], (e) IST-CRF [32], (f) GBF-CD [29], (g) GLSS [30], (h) CANet [47], (i) CACD [20], (j) SR-GCAE [50], (k) BAACL [42], and (l) proposed AEKAN.

well in the evaluation indicators of all datasets, except that in dataset #4, the OA indicator is slightly lower than the highest value, with only a difference of 0.001, which is still the second-best. In other evaluation metrics on all datasets, our method achieves the best performance. Especially in dataset #2, our K_a is 0.046 higher than the next highest value, and even in the more complex dataset #5, our method still achieves the K_a of 0.658. Therefore, significant improvements in average K_a and F_1 of 0.074 and 0.07 can be obtained compared to the second-best method AGSCC [15]. Therefore, in addition to the above visual comparison results, quantitative accuracy comparison also supports the conclusions of the

visual comparison. Overall, combining the above quantitative evaluation results with the visual analysis results can fully demonstrate the effectiveness and superiority of the proposed AEKAN in the MCD task.

V. DISCUSSION

In this section, four experiments are conducted to further analyze and discuss the effectiveness of the proposed AEKAN. First, the effect of the proposed hierarchical commonality loss is investigated. Second, we visualize the features of different stages in the proposed AEKAN by t-SNE [58] to intuitively analyze the performance of commonality features. Finally, the

TABLE III
COMPARISONS BETWEEN THE PROPOSED METHOD AND OTHER RELATED SOTA METHODS FOR DATASETS #1–#5. THE BEST RESULTS AND SECOND-BEST ACCURACY ARE PRESENTED IN BOLD RED AND BLUE, RESPECTIVELY

Methods	Dataset #1			Dataset #2			Dataset #3			Dataset #4			Dataset #5			Average		
	OA	Ka	F1															
IRG-MCs [31]	0.983	0.794	0.803	0.936	0.704	0.740	0.950	0.606	0.633	0.959	0.490	0.512	0.875	0.349	0.403	0.941	0.589	0.618
GIR-MRF [33]	0.981	0.798	0.807	0.937	0.734	0.770	0.957	0.711	0.734	0.960	0.491	0.512	0.873	0.447	0.519	0.945	0.636	0.669
SCASC [14]	0.979	0.741	0.751	0.950	0.776	0.804	0.980	0.828	0.838	0.960	0.479	0.499	0.872	0.327	0.381	0.948	0.630	0.655
AGSCC [15]	0.983	0.772	0.781	0.955	0.791	0.817	0.979	0.809	0.819	0.960	0.486	0.506	0.895	0.489	0.542	0.954	0.669	0.693
IST-CRF [32]	0.981	0.801	0.811	0.920	0.649	0.695	0.935	0.585	0.618	0.958	0.527	0.549	0.899	0.614	0.674	0.939	0.635	0.670
GBF-CD [29]	0.720	0.161	0.229	0.654	0.197	0.345	0.586	0.072	0.176	0.870	0.259	0.309	0.690	0.114	0.289	0.704	0.160	0.270
GLSS [30]	0.854	0.200	0.256	0.850	0.490	0.571	0.642	0.141	0.236	0.932	0.433	0.466	0.459	-0.100	0.155	0.747	0.233	0.337
CANet [47]	0.954	0.414	0.437	0.844	0.225	0.313	0.854	-0.026	0.511	0.935	0.359	0.392	0.838	0.199	0.277	0.885	0.234	0.386
X-Net [16]	0.954	0.595	0.618	0.909	0.637	0.688	-	-	-	0.907	0.354	0.395	0.830	0.229	0.322	0.900	0.454	0.506
ACE-Net [16]	0.956	0.613	0.635	0.928	0.659	0.701	-	-	-	0.923	0.402	0.438	0.842	0.294	0.382	0.912	0.492	0.539
CACD [20]	0.978	0.738	0.750	0.798	0.417	0.516	0.906	0.190	0.240	0.743	-0.076	0.002	0.829	0.029	0.085	0.851	0.260	0.319
SR-GCAE [50]	0.982	0.761	0.770	0.884	0.586	0.649	0.961	0.660	0.681	0.885	0.325	0.370	0.874	0.460	0.531	0.917	0.558	0.600
BAACL [42]	0.978	0.686	0.697	0.810	0.412	0.510	0.923	0.421	0.460	0.842	0.204	0.260	0.789	0.254	0.380	0.868	0.395	0.462
Proposed AEKAN	0.987	0.831	0.838	0.964	0.837	0.858	0.983	0.839	0.847	0.959	0.551	0.572	0.921	0.658	0.702	0.963	0.743	0.763

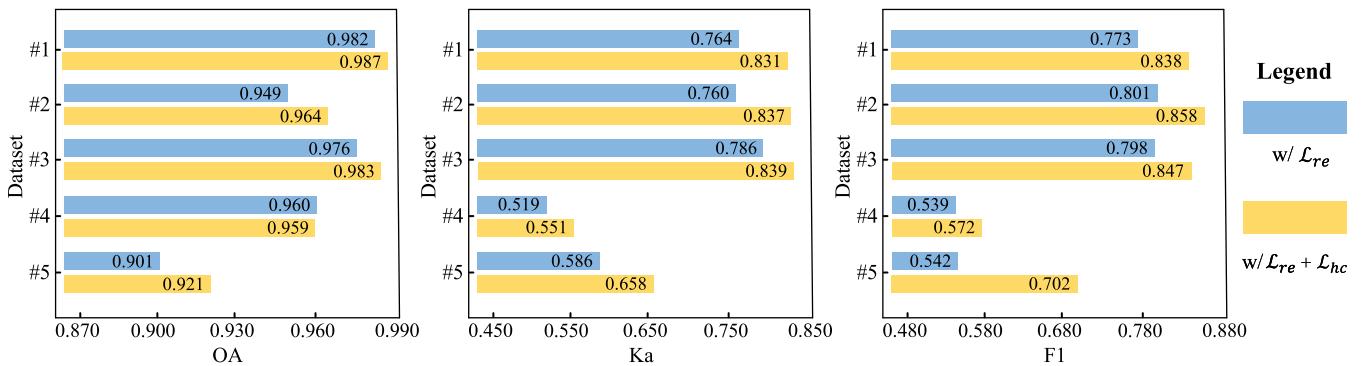


Fig. 6. Results of ablation studies on MCD datasets #1–#5 for the proposed hierarchical commonality loss \mathcal{L}_{hc} using three evaluation metrics (OA, Ka, and F1).

sensitivity analysis of two parameters (\mathcal{P}_n and \mathcal{P}_c) involved in the proposed AEKAN is tested and discussed, respectively.

A. Hierarchical Commonality Loss Analysis

In the proposed AEKAN, we propose a hierarchical commonality loss \mathcal{L}_{hc} to explicitly guide the Siamese KAN encoder in the proposed AEKAN to obtain and enhance the latent commonality features between HRSIs. To verify the effectiveness of the hierarchical commonality loss \mathcal{L}_{hc} , we conducted ablation experiments on datasets #1–#5 to train the model, including w/\mathcal{L}_{re} and $w/\mathcal{L}_{re} + \mathcal{L}_{hc}$. Based on this, the ablation studies of hierarchical commonality loss were performed, and the results are shown in Fig. 6. The overall results reflect that the model after adding hierarchical commonality loss \mathcal{L}_{hc} shows higher detection accuracy in the three evaluation indicators of OA, Ka, and F1. For instance, for dataset #1, the Ka increased from 0.764 to 0.831 (an increase of 0.067), and the F1 improved from 0.773 to 0.838 (an increase of 0.065); for dataset #5, the Ka and F1 boost 0.072 and 0.16, respectively. These results reveal that introducing the hierarchical commonality loss \mathcal{L}_{hc} has significant advantages in improving the detection accuracy of the proposed AEKAN. Therefore, the hierarchical commonality loss \mathcal{L}_{hc} can encourage our Siamese KAN encoder to better learn and extract the commonality features between HRSIs, which can be effectively used for MCD by directly comparing these commonality features.

B. Feature Visualization Analysis

In this section, to deeply explore the role of hierarchical commonality loss \mathcal{L}_{hc} in the training process, we adopted t-SNE dimensionality reduction technology to visualize the features of different stages in the proposed AEKAN. Specifically, it includes the following four contents: First, we visualized the features of the original HRSIs in the two modalities without training to display the feature distribution of the original HRSIs. Subsequently, we performed inference on the trained Siamese KAN encoder and exploited t-SNE dimensionality reduction technology to visualize the commonality feature distribution of the HRSIs output by the first and second KAN layers to present the changes in the commonality feature distribution after training. Finally, we also visualized the feature distribution of DIs and BCIs.

Fig. 7 provides all visualization results. By observing Fig. 7(a), we can find the feature distribution of the original HRSIs exhibits the difference between the two modalities obviously. Specifically, for datasets #1–#3, the blue and red dots are clearly separated due to the significant imaging differences between optical images and SAR, resulting in significantly different spectra. In contrast, for the HRSIs in datasets #4 and #5, the differences are not significant because their bi-temporal images are both optical images. In addition,

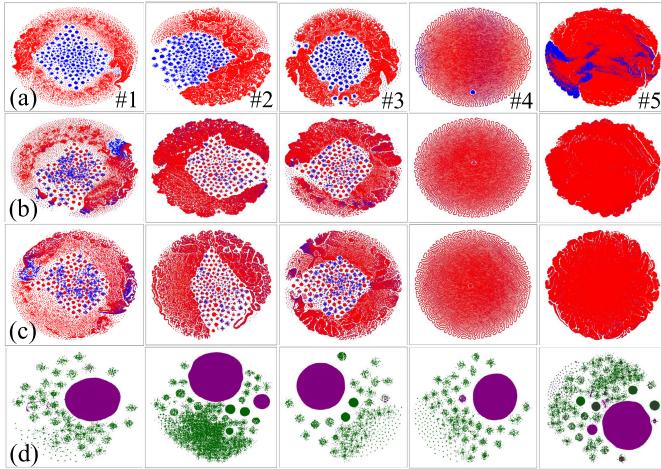


Fig. 7. Feature visualization analysis via t-SNE for datasets #1–#5. (a) Original HRSIs. (b) Commonality features of the HRSIs output by the first KAN layer in the Siamese KAN encoder. (c) Commonality features of the HRSIs output by the second KAN layer in the Siamese KAN encoder. (d) DIs. Note: Red and blue dots represent the pixels within pre and post-event images; green and purple dots denote unchanged and changed pixels, respectively.

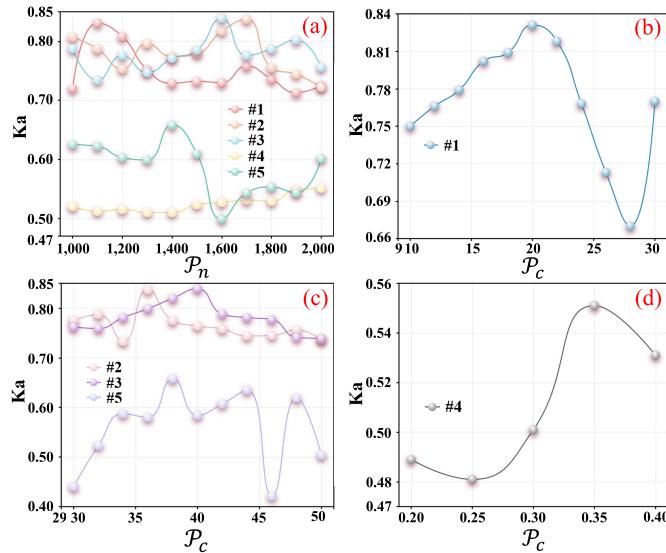


Fig. 8. Parameter sensitivity analysis of the parameters P_n and P_c in the proposed superpixel-based AEKAN on datasets #1–#5. (a) Results of the parameter P_n . (b)–(d) Results of the parameter P_c .

Fig. 7(b) and (c) demonstrates that the feature distribution of the first and second KAN layers of the encoder, for datasets #1–#3, the blue and red points in some areas show obvious overlap, indicating that the KAN encoder captures effectively commonality features. Overall, before training, the differences between modalities appear as obvious separations in the feature space, while after training, the overlapping areas between modalities increase significantly as the Siamese KAN encoder learns the commonality features. This also intuitively reveals the effectiveness of the hierarchical commonality loss \mathcal{L}_{hc} . Furthermore, Fig. 7(d) presents that the differences between HRSIs in MCD can be effectively measured by the distance between the commonality features of HRSIs. In summary, this experiment has demonstrated that the commonality features in the unchanged areas are significantly more common than those in the changed areas, which further proves that the commonality features can be utilized to detect changes between HRSIs to realize MCD.

C. Parameter Analysis for \mathcal{P}_n

In the proposed AEKAN, the parameter \mathcal{P}_n is the segmentation granularity in the SLIC algorithm by adjusting the number of superpixels. Consequently, different segmentation granularities may affect the performance of the proposed AEKAN. To this end, we analyze the effect of the parameter \mathcal{P}_n in the proposed AEKAN on the performance (in Ka). Herein, we set the value of \mathcal{P}_n to a value every 100 in the range of [1000, 2000] to test the performance of different \mathcal{P}_n on datasets #1–#5.

Based on this setting, we plotted the Ka accuracy as a function of the number of superpixels to visually observe the performance of each dataset under different parameters \mathcal{P}_n , as shown in Fig. 8(a). Observing Fig. 8(a), it can be seen that the performance of the proposed AEKAN on datasets #1–#3 is relatively stable, and Ka fluctuates almost between 0.7 and 0.85. For dataset #4, Ka has no major changes and is stable between 0.5 and 0.555, showing a gradually increasing trend. For the dataset #5, the proposed AEKAN fluctuates greatly and reaches a peak at 1400 and a trough at 1600. In general, the sensitivity of the parameter \mathcal{P}_n on the proposed AEKAN is also different due to different types and scales of objects covered by the dataset. That is because the parameter \mathcal{P}_n determines the number and scale of superpixels. Concretely, a larger value of the parameter \mathcal{P}_n means more superpixels, which makes the scale covered by each superpixel smaller. This may lead to insufficient information and difficulty in effectively extracting commonality features, resulting in missed detection or false detection. On the contrary, a smaller parameter \mathcal{P}_n means fewer superpixels, which may make the scale of each superpixel too large. It may introduce other land cover objects due to over-segmentation, leading to false detection. On the whole, the proposed AEKAN needs to select appropriate parameters \mathcal{P}_n for different MCD datasets to achieve the best performance.

D. Parameter Analysis for \mathcal{P}_c

In addition to the parameter \mathcal{P}_n , in the proposed AEKAN, the parameter \mathcal{P}_c can affect the superpixel shape and boundary smoothness in the SLIC algorithm. To analyze the influence of the parameter \mathcal{P}_c on the performance of the proposed AEKAN, we study the relationship between different parameters \mathcal{P}_c and Ka on datasets #1–#5. In the experiments, since the land cover types and ranges of different MCD datasets are different we set different values for different datasets #1–#5, that is: for dataset #1, the value of \mathcal{P}_c is set to every 2 values in the range of [10, 30]; for datasets #2, #3, and #5, the value of \mathcal{P}_c is set to every 2 values in the range of [30, 50]; For dataset #4, the value of \mathcal{P}_c is set to [0.2, 0.25, 0.3, 0.35, 0.4].

According to the above settings, the relationship curves of the \mathcal{P}_c and Ka on datasets #1–#5 are presented in Fig. 8(b)–(d): dataset #1, datasets #2, #3, and #5, and dataset #4, respectively. From the curve analysis, it can be seen that the fluctuations of datasets #2 and #3 are relatively small, while the fluctuations of the other three datasets #1, #4, and #5 are relatively large. For example, for dataset #5, the Ka performance of the parameter \mathcal{P}_c fluctuates between 0.4, and 0.7, but in most cases, it is concentrated around 0.6. This is

TABLE IV
RESULTS OF THE PROPOSED METHOD BASED ON MLP AND KAN LAYERS ON DATASETS #1–#5, RESPECTIVELY. THE BEST RESULTS ACCURACY ARE PRESENTED IN BOLD RED

Methods	Dataset #1			Dataset #2			Dataset #3			Dataset #4			Dataset #5		
	OA	KC	F1												
w/ MLP	0.984	0.790	0.798	0.945	0.778	0.809	0.975	0.776	0.789	0.956	0.404	0.426	0.913	0.580	0.624
w/ KAN	0.987	0.831	0.834	0.964	0.837	0.858	0.983	0.839	0.847	0.959	0.551	0.572	0.921	0.658	0.702

because higher parameter \mathcal{P}_c produces more regular square superpixels, while lower \mathcal{P}_c allows superpixels to adapt more freely to boundaries and details in the image. Hence, it is necessary to choose a suitable parameter \mathcal{P}_c to reach the best MCD results.

E. Ablation Analysis of KAN Layer

To verify the effect of the KAN layer, we construct an ablation experiment in this section. To this end, we designed a comparative experiment to replace the pure KAN layer within the proposed AEKAN with a multilayer perceptron (MLP) with the same settings. The results are listed in Table IV. According to the quantitative results in the table, the KAN network performs significantly better than MLP on all data sets in the MCD task. For example, for datasets #2–#4, KAN's K_a values are 5.9%, 6.3%, and 14.7% higher than MLP, respectively. Furthermore, the F_1 value of KAN is also 7.8% higher on dataset #5. These results fully demonstrate the effectiveness and advantages of the KAN in the MCD task.

VI. CONCLUSION AND FUTURE WORKS

In this study, we proposed a novel superpixel-based AEKAN for unsupervised MCD. In the proposed method, we explore the effectiveness of KAN for feature representation in MCD. Concretely, we can extract the commonality features of HRSIs by the proposed Siamese KAN encoder to overcome the limitation that it cannot capture the changes by directly comparing HRSIs. Moreover, we devised a hierarchical commonality loss to explicitly guide the Siamese KAN encoder in the proposed AEKAN to capture commonality features effectively. Based on the extracted commonality features, a high-quality DI can be generated for MCD. Extensive comparative experiments have demonstrated the superiority of the proposed AEKAN on the MCD task. In addition, sufficient ablation studies have demonstrated the existence of commonality features between HRSIs and their ability to identify changes between HRSIs. Although the proposed AEKAN has exhibited certain capabilities and advantages in MCD, it still has the following limitations that need to be further improved. That is, the stability of the proposed AEKAN is still insufficient due to the lack of any prior supervision information. Moreover, the effectiveness of the proposed method in more sensors, large scales, and complex scenes also needs to be further verified. In future work, we will continue to explore the adaptability and robustness of the proposed AEKAN while promoting its stability.

ACKNOWLEDGMENT

The authors are grateful to the editor-in-chief, associate editor, and reviewers for their insightful comments and suggestions. They would like to appreciate the researchers of the opening datasets and codes for their contributions.

REFERENCES

- [1] W. Shi, M. Zhang, R. Zhang, S. Chen, and Z. Zhan, “Change detection based on artificial intelligence: State-of-the-art and challenges,” *Remote Sens.*, vol. 12, no. 10, p. 1688, May 2020.
- [2] Z. Lv, T. Liu, J. A. Benediktsson, and N. Falco, “Land cover change detection techniques: Very-high-resolution optical images: A review,” *IEEE Geosci. Remote Sens. Mag.*, vol. 10, no. 1, pp. 44–63, Mar. 2022.
- [3] Z. Lv et al., “Land cover change detection with heterogeneous remote sensing images: Review, progress, and perspective,” *Proc. IEEE*, vol. 110, no. 12, pp. 1976–1991, Dec. 2022.
- [4] K. M. Bergen, D. G. Brown, J. F. Rutherford, and E. J. Gustafson, “Change detection with heterogeneous data using ecoregional stratification, statistical summaries and a land allocation algorithm,” *Remote Sens. Environ.*, vol. 97, no. 4, pp. 434–446, Sep. 2005.
- [5] Y. Sun, L. Lei, X. Li, X. Tan, and G. Kuang, “Patch similarity graph matrix-based unsupervised remote sensing change detection with homogeneous and heterogeneous sensors,” *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 6, pp. 4841–4861, Jun. 2021.
- [6] Y. Sun, L. Lei, X. Li, H. Sun, and G. Kuang, “Nonlocal patch similarity based heterogeneous remote sensing change detection,” *Pattern Recognit.*, vol. 109, Jan. 2021, Art. no. 107598.
- [7] L. Lei, Y. Sun, and G. Kuang, “Adaptive local structure consistency-based heterogeneous remote sensing change detection,” *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022.
- [8] Y. Sun, L. Lei, X. Li, X. Tan, and G. Kuang, “Structure consistency-based graph for unsupervised change detection with homogeneous and heterogeneous remote sensing images,” *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 4700221.
- [9] R. Touati and M. Mignotte, “An energy-based model encoding nonlocal pairwise pixel interactions for multisensor change detection,” *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 2, pp. 1046–1058, Feb. 2018.
- [10] R. Touati, M. Mignotte, and M. Dahmane, “A new change detector in heterogeneous remote sensing imagery,” in *Proc. 7th Int. Conf. Image Process. Theory, Tools Appl. (IPTA)*, Nov. 2017, pp. 1–6.
- [11] R. Touati, M. Mignotte, and M. Dahmane, “Multimodal change detection in remote sensing images using an unsupervised pixel pairwise-based Markov random field model,” *IEEE Trans. Image Process.*, vol. 29, pp. 757–767, 2020.
- [12] Z. Liu, G. Li, G. Mercier, Y. He, and Q. Pan, “Change detection in heterogeneous remote sensing images via homogeneous pixel transformation,” *IEEE Trans. Image Process.*, vol. 27, no. 4, pp. 1822–1834, Apr. 2018.
- [13] L. T. Luppino, F. M. Bianchi, G. Moser, and S. N. Anfinsen, “Unsupervised image regression for heterogeneous change detection,” *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 12, pp. 9960–9975, Dec. 2019.
- [14] Y. Sun, L. Lei, D. Guan, M. Li, and G. Kuang, “Sparse-constrained adaptive structure consistency-based unsupervised image regression for heterogeneous remote-sensing change detection,” *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 4405814.
- [15] Y. Sun, L. Lei, D. Guan, J. Wu, G. Kuang, and L. Liu, “Image regression with structure cycle consistency for heterogeneous change detection,” *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 2, pp. 1613–1627, Feb. 2022.

- [16] L. T. Luppino et al., "Deep image translation with an affinity-based change prior for unsupervised multimodal change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 4700422.
- [17] Z.-G. Liu, Z.-W. Zhang, Q. Pan, and L.-B. Ning, "Unsupervised change detection from heterogeneous data based on image translation," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 4403413.
- [18] X. Li, Z. Du, Y. Huang, and Z. Tan, "A deep translation (GAN) based change detection network for optical and SAR remote sensing images," *ISPRS J. Photogramm. Remote Sens.*, vol. 179, pp. 14–34, Sep. 2021.
- [19] J.-J. Wang, N. Dobigeon, M. Chabert, D.-C. Wang, T.-Z. Huang, and J. Huang, "CD-GAN: A robust fusion-based generative adversarial network for unsupervised remote sensing change detection with heterogeneous sensors," *Inf. Fusion*, vol. 107, Jul. 2024, Art. no. 102313.
- [20] Y. Wu, J. Li, Y. Yuan, A. K. Qin, Q.-G. Miao, and M.-G. Gong, "Commonality autoencoder: Learning common features for change detection from heterogeneous images," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 9, pp. 4257–4270, Sep. 2022.
- [21] G. Mercier, G. Moser, and S. B. Serpico, "Conditional copulas for change detection in heterogeneous remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 5, pp. 1428–1441, May 2008.
- [22] J. Prendes, M. Chabert, F. Pascal, A. Giros, and J.-Y. Tourneret, "A new multivariate statistical model for change detection in images acquired by homogeneous and heterogeneous sensors," *IEEE Trans. Image Process.*, vol. 24, no. 3, pp. 799–812, Mar. 2015.
- [23] Z.-G. Liu, G. Mercier, J. Dezert, and Q. Pan, "Change detection in heterogeneous remote sensing images based on multidimensional evidential reasoning," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 1, pp. 168–172, Jan. 2014.
- [24] M. Volpi, F. de Morsier, G. Camps-Valls, M. Kanevski, and D. Tuia, "Multi-sensor change detection based on nonlinear canonical correlations," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2013, pp. 1944–1947.
- [25] R. Touati, M. Mignotte, and M. Dahmane, "Change detection in heterogeneous remote sensing images based on an imaging modality-invariant MDS representation," in *Proc. IEEE Int. Conf. Image Process.*, Oct. 2018, pp. 3998–4002.
- [26] M. Mignotte, "A fractal projection and Markovian segmentation-based approach for multimodal change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 11, pp. 8046–8058, Nov. 2020.
- [27] M. Gong, P. Zhang, L. Su, and J. Liu, "Coupled dictionary learning for change detection from multisource data," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 12, pp. 7077–7091, Dec. 2016.
- [28] V. Ferraris, N. Dobigeon, Y. Cavalcanti, T. Oberlin, and M. Chabert, "Coupled dictionary learning for unsupervised change detection between multimodal remote sensing images," *Comput. Vis. Image Understand.*, vol. 189, Dec. 2019, Art. no. 102817.
- [29] D. A. Jimenez-Sierra, H. D. Benítez-Restrepo, H. D. Vargas-Cardona, and J. Chanussot, "Graph-based data fusion applied to: Change detection and biomass estimation in rice crops," *Remote Sens.*, vol. 12, no. 17, p. 2683, Aug. 2020.
- [30] D. A. Jimenez-Sierra, D. A. Quintero-Olaya, J. C. Alvear-Muñoz, H. D. Benítez-Restrepo, J. F. Florez-Ospina, and J. Chanussot, "Graph learning based on signal smoothness representation for homogeneous and heterogeneous change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 4410416.
- [31] Y. Sun, L. Lei, D. Guan, and G. Kuang, "Iterative robust graph for unsupervised change detection of heterogeneous remote sensing images," *IEEE Trans. Image Process.*, vol. 30, pp. 6277–6291, 2021.
- [32] Y. Sun, L. Lei, D. Guan, J. Wu, and G. Kuang, "Iterative structure transformation and conditional random field based method for unsupervised multimodal change detection," *Pattern Recognit.*, vol. 131, Nov. 2022, Art. no. 108845.
- [33] Y. Sun, L. Lei, X. Tan, D. Guan, J. Wu, and G. Kuang, "Structured graph based image regression for unsupervised multimodal change detection," *ISPRS J. Photogramm. Remote Sens.*, vol. 185, pp. 16–31, Mar. 2022.
- [34] L. Zhu, W. Sun, D. Fan, H. Xing, and X. Liu, "Unsupervised spatial self-similarity difference-based change detection method for multi-source heterogeneous images," *Pattern Recognit.*, vol. 149, May 2024, Art. no. 110237.
- [35] T. Han, Y. Tang, B. Zou, and H. Feng, "Unsupervised multimodal change detection based on adaptive optimization of structured graph," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 126, Feb. 2024, Art. no. 103630.
- [36] G. Cheng et al., "Change detection methods for remote sensing in the last decade: A comprehensive review," *Remote Sens.*, vol. 16, no. 13, p. 2355, Jun. 2024.
- [37] R. Shao, C. Du, H. Chen, and J. Li, "SUNet: Change detection for heterogeneous remote sensing images from satellite and UAV using a dual-channel fully convolution network," *Remote Sens.*, vol. 13, no. 18, p. 3750, Sep. 2021.
- [38] Z. Lv, J. Liu, W. Sun, T. Lei, J. A. Benediktsson, and X. Jia, "Hierarchical attention feature fusion-based network for land cover change detection with homogeneous and heterogeneous remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 4411115.
- [39] H. Li, M. Gong, M. Zhang, and Y. Wu, "Spatially self-paced convolutional networks for change detection in heterogeneous images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 4966–4979, 2021.
- [40] J. Shi, T. Wu, A. K. Qin, Y. Lei, and G. Jeon, "Self-guided autoencoders for unsupervised change detection in heterogeneous remote sensing images," *IEEE Trans. Artif. Intell.*, vol. 5, no. 6, pp. 2458–2471, Jun. 2024.
- [41] D. Wang, F. Zhao, H. Yi, Y. Li, and X. Chen, "An unsupervised heterogeneous change detection method based on image translation network and post-processing algorithm," *Int. J. Digit. Earth*, vol. 15, no. 1, pp. 1056–1080, Dec. 2022.
- [42] M. Jia, C. Zhang, Z. Zhao, and L. Wang, "Bipartite graph attention autoencoders for unsupervised change detection using VHR remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5626215.
- [43] W. Cheng, Y. Feng, L. Song, and X. Wang, "DMF2Net: Dynamic multi-level feature fusion network for heterogeneous remote sensing image change detection," *Knowl.-Based Syst.*, vol. 300, Sep. 2024, Art. no. 112159.
- [44] T. Zhan, M. Gong, X. Jiang, and S. Li, "Log-based transformation feature learning for change detection in heterogeneous images," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 9, pp. 1352–1356, Sep. 2018.
- [45] M. Jia, Q. Zhao, and X. Lu, "Bipartite adversarial autoencoders for unsupervised heterogeneous remote sensing image change detection," (in chinese), *Int. J. Image Graph. Int.*, vol. 29, no. 8, pp. 2188–2204, 2024.
- [46] J. Liu, M. Gong, K. Qin, and P. Zhang, "A deep convolutional coupling network for change detection based on heterogeneous optical and radar images," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 3, pp. 545–559, Mar. 2018.
- [47] X. Niu, M. Gong, T. Zhan, and Y. Yang, "A conditional adversarial network for change detection in heterogeneous images," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 1, pp. 45–49, Jan. 2019.
- [48] W. Zhao, Z. Wang, M. Gong, and J. Liu, "Discriminative feature learning for unsupervised change detection in heterogeneous images based on a coupled neural network," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 12, pp. 7066–7080, Dec. 2017.
- [49] Q. Liu, K. Ren, X. Meng, and F. Shao, "Domain adaptive cross reconstruction for change detection of heterogeneous remote sensing images via a feedback guidance mechanism," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 4507216.
- [50] H. Chen, N. Yokoya, C. Wu, and B. Du, "Unsupervised multimodal change detection based on structural relationship graph representation learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, Dec. 2022, Art. no. 5635318.
- [51] H. Chen, N. Yokoya, and M. Chini, "Fourier domain structural relationship analysis for unsupervised multimodal change detection," *ISPRS J. Photogramm. Remote Sens.*, vol. 198, pp. 99–114, Apr. 2023.
- [52] Z. Liu et al., "KAN: Kolmogorov–Arnold networks," 2024, *arXiv:2404.19756*.
- [53] M. Cheon, "Kolmogorov–Arnold network for satellite image classification in remote sensing," 2024, *arXiv:2406.00600*.
- [54] A. Jamali, S. K. Roy, D. Hong, B. Lu, and P. Ghamisi, "How to learn more? Exploring Kolmogorov–Arnold networks for hyperspectral image classification," 2024, *arXiv:2406.15719*.
- [55] V. Lobanov, N. Firsov, E. Myasnikov, R. Khabibullin, and A. Nikonorov, "HyperKAN: Kolmogorov–Arnold networks make hyperspectral image classifiers smarter," 2024, *arXiv:2407.05278*.
- [56] Y. Wang et al., "SpectralKAN: Kolmogorov–Arnold network for hyperspectral images change detection," 2024, *arXiv:2407.00949*.
- [57] N. Otsu, "A threshold selection method from gray-level histograms," *Automatica*, vol. 11, nos. 285–296, pp. 23–27, 1975.
- [58] L. Van der Maaten and G. Hinton, "Visualizing data using t-SNE," *J. Mach. Learn. Res.*, vol. 9, no. 11, pp. 2579–2605, Jan. 2008.



Tongfei Liu (Member, IEEE) received the master's degree from the School of Computer Science and Engineering, Xi'an University of Technology, Xi'an, China, in 2020, and the Ph.D. degree from the School of Electronic Engineering, Xidian University, Xi'an, in 2023.

He is currently a Lecturer with the School of Electronic Information and Artificial Intelligence, Shaanxi University of Science and Technology, Xi'an. His research interests include deep learning, spatial-spectral feature extraction, pattern recognition, building extraction, and land cover change detection and classification, through very-high resolution (VHR) remote sensing images (including satellite and aerial images).



Weichuan Zhang (Member, IEEE) received the M.S. degree in signal and information processing from Southwest Jiaotong University, Chengdu, China, in 2006, and the Ph.D. degree in signal and information processing from the National Laboratory of Radar Signal Processing, Xidian University, Xi'an, China, in 2012.

He is currently a Professor with the School of Electronic Information and Artificial Intelligence, Shaanxi University of Science and Technology, Xi'an, and a Principal Research Fellow with Griffith University, Nathan, QLD, Australia. His research interests include computer vision, image analysis, and pattern recognition.



Jianjian Xu received the B.S. degree from the School of Information Engineering, Xi'an Mingde Institute of Technology, Xi'an, China, in 2022. He is currently pursuing the master's degree with the School of Electronic Information and Artificial Intelligence, Shaanxi University of Science and Technology, Xi'an.

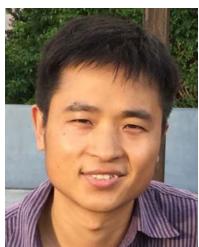
His interests include deep learning and multimodal change detection.



Zhiyong Lv (Senior Member, IEEE) received the M.S. and Ph.D. degrees from the School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, China, in 2008 and 2014, respectively.

He was an Engineer in surveying with The First Institute of Photogrammetry and Remote Sensing, Xi'an, China, from 2008 to 2011. He is currently working with the School of Computer Science and Engineering, Xi'an University of Technology, Xi'an. His research interests include multi hyperspectral

and high-resolution remotely sensed image processing, spatial feature extraction, neural networks, pattern recognition, deep learning, and remote sensing applications.



Tao Lei (Senior Member, IEEE) received the Ph.D. degree in information and communication engineering from Northwestern Polytechnical University, Xi'an, China, in 2011.

From 2012 to 2014, he was a Post-Doctoral Research Fellow with the School of Electronics and Information, Northwestern Polytechnical University. From 2015 to 2016, he was a Visiting Scholar with the Quantum Computation and Intelligent Systems Group, University of Technology Sydney, Sydney, NSW, Australia. He is currently a Professor with the School of Electronic Information and Artificial Intelligence, Shaanxi University of Science and Technology, Xi'an. He has authored or co-authored more than 80 research articles, including IEEE TRANSACTIONS ON IMAGE PROCESSING (TIP), IEEE TRANSACTIONS ON FUZZY SYSTEMS (TFS), and IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING (TGRS). His research interests include image processing, pattern recognition, and machine learning.



Yingbo Wang (Member, IEEE) received the Ph.D. degree from the School of Optics and Photonics, Beijing Institute of Technology, Beijing, China, in 2012.

He is currently a Lecturer with the School of Electronic Information and Artificial Intelligence, Shaanxi University of Science and Technology, Xi'an, China. His current research interests include image restoration, computer vision, and deep learning.

Maoguo Gong (Fellow, IEEE) received the B.S. degree (Hons.) in electronic engineering and the Ph.D. degree in electronic science and technology from Xidian University, Xi'an, China, in 2003 and 2009, respectively.

Since 2006, he has been a Teacher with Xidian University. He was promoted to an Associate Professor and a Full Professor, in 2008 and 2010, respectively, both with exceptive admission. He is leading or has completed over 20 projects as the Principle Investigator, funded by the National Natural Science Foundation of China and the National Key Research and Development Program of China, and others. He has authored or co-authored over 100 articles in journals and conferences, and holds over 20 granted patents as the First Inventor. His research interests are broadly in the area of computational intelligence, with applications to optimization, learning, data mining, and image understanding.

Dr. Gong is the Executive Committee Member of the Chinese Association for Artificial Intelligence and a Senior Member of the Chinese Computer Federation. He was a recipient of the Prestigious National Program for Support of the Leading Innovative Talents (selected by the Central Organization Department of China), the Leading Innovative Talent in Science and Technology (selected by the Ministry of Science and Technology of China), the Excellent Young Scientist Foundation (selected by the National Natural Science Foundation of China), the New Century Excellent Talent in University (selected by the Ministry of Education of China), the Young Teacher Award by the Fok Ying Tung Education Foundation, and the National Natural Science Award of China. He is an Associate Editor or an Editorial Board Member for over five journals including IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION and IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS.



Xiaogang Du received the B.S., M.S., and Ph.D. degrees from Lanzhou Jiaotong University, Lanzhou, China, in 2007, 2010, and 2017, respectively.

He is currently an Associate Professor with the School of Electronic Information and Artificial Intelligence, Shaanxi University of Science and Technology, Xi'an, China. His research interests include computer vision, machine learning, and deep learning.

