

Detecting changes without comparing images: Rules induced change detection in heterogeneous remote sensing images

Yuli Sun ^{a,b}, Lin Lei ^a, Zhang Li ^b, Gangyao Kuang ^a, Qifeng Yu ^b

^a College of Electronic science and Technology, National University of Defense Technology, Changsha 410073, China

^b College of Aerospace Science and Engineering, National University of Defense Technology, Changsha 410073, China



ARTICLE INFO

Keywords:

Change detection
Heterogeneous data
Multimodal
Energy based model
Remote sensing images

ABSTRACT

Heterogeneous change detection (HCD) is crucial for monitoring surface changes using various remote sensing data, especially in disaster emergency response and environmental monitoring. To facilitate the comparability of heterogeneous images, previous methods are devoted to designing various complex transformation functions to transfer heterogeneous images into a common domain for comparison. As a result, the performance of HCD is constrained by the accuracy and robustness of these transformation functions. Unlike existing comparison-based HCD methods that rely on complex transformations and feature alignments between heterogeneous images, this paper proposes an unsupervised rules-induced energy model (RIEM) that detects changes by independently analyzing intra-image relationships, without explicitly comparing the heterogeneous images. This frees HCD from the complicated and challenging transformations and interactions between heterogeneous images. Specifically, we first establish the connections between the class relationships (same/different) and change labels (changed/unchanged) of pairwise superpixels, and then derive six rules for determining the change label of each superpixel, which enables detecting changes by considering only the intra-image relationships within each image, without inter-image comparisons. Then, we build an energy-based model to release the ability of rules to identify changes, which implements four types of energy loss functions. Remarkably, since the rules used in the energy model are derived based on the nature of change detection problem, the proposed RIEM is highly robust to imaging conditions. Extensive experiments on seven datasets demonstrate the efficacy of RIEM in detecting changes from heterogeneous images. The code is released at <https://github.com/yulisin/RIEM>.

1. Introduction

1.1. Background

Heterogeneous change detection (HCD) refers to the technique of extracting land cover change information by analyzing multitemporal remote sensing (RS) images captured over the same geographical region but under heterogeneous conditions with different sensors (e.g., SAR and optical sensors) (Lv et al., 2022b; Mercier et al., 2008). HCD can be viewed as an extension of the homogeneous change detection, in the latter multitemporal RS images come from the same sensor with similar imaging conditions (Wu et al., 2023; Zheng et al., 2024; Chen et al., 2023a).

Compared to the conventional homogeneous change detection, HCD offers three main advantages: (1) HCD allows for the rapid acquisition of change information by utilizing any available pre-change and post-change images from different sources when obtaining homogeneous images is impossible due to imaging constraints such as lighting

and weather conditions (Touati, 2019; Luppino, 2020). (2) With the usage of heterogeneous images, HCD can enhance the temporal resolution of change analysis when obtaining homogeneous images within a short time window is infeasible due to satellite revisit limitations (Sun et al., 2021b; McGregor et al., 2024). (3) HCD can fully leverage early RS image acquired by outdated sensors, thereby expanding the temporal horizon for time-series monitoring and analysis (Sun et al., 2022). Based on these advantages, HCD has been widely applied in various fields, particularly in emergency response scenarios such as floods, fires, earthquakes, landslides, explosions (Lv et al., 2022b; Zhao et al., 2024). Moreover, with the rapid advancement of space and imaging technologies, an increasing variety of new sensors are being deployed (Zheng et al., 2023; Hong et al., 2024), while older satellites have a limited lifespan and eventually cease operation. Given its ability to overcome sensor-specific limitations and adverse imaging conditions, HCD is gradually becoming an attractive topic in the field of remote sensing (Yu et al., 2023; Shi et al., 2024).

* Corresponding author.

E-mail address: sunyuli@mail.ustc.edu.cn (Y. Sun).

1.2. Related work

Despite its significant application prospects, HCD also faces substantial challenges compared to homogeneous change detection. Because of the different imaging mechanisms, imaging conditions and imaging modes among different sensors, heterogeneous images exhibit dramatical variations in the representation of the same land features, which makes it challenging to identify changes by comparing the multitemporal images directly as done in the homogeneous change detection (Wu et al., 2024b; Jia et al., 2022b; Chen et al., 2024b). So, the primary task of HCD is to establish the imaging-independent correlations between heterogeneous images and make them “comparable” (Sun et al., 2021b). To achieve this goal, researchers have proposed different methods for HCD and they mostly use the paradigm of $DI = \mathcal{M}_1(X) \ominus \mathcal{M}_2(Y)$, where DI denotes the difference image that measures the change score for each pixel, \mathcal{M}_1 and \mathcal{M}_2 denote the transformation functions that transform pre-change image X and post-change image Y into the same domain respectively, and \ominus denotes the pixelwise difference operator (Liu et al., 2018a). Therefore, previous HCD methods have worked on exploiting better operators of \mathcal{M}_1 , \mathcal{M}_2 and \ominus to improve the detection performance.

Depending on the transformed common domain, these comparison-based methods further contain methods that compare images within the same category domain, within the same feature domain, and within the same image domain. For example, (i) the post-classification comparison methods first classify the pre- and post-change images into a common land cover category, and then compare the classification results to extract change information, such as the Bayesian classification method (Camps-Valls et al., 2008), compound classification method (Wan et al., 2019), and evidence reasoning based method (Liu et al., 2014). (ii) The feature transformation based methods utilize some imaging-invariant assumptions or deep Siamese/pseudo-Siamese network to extract shared features from heterogeneous images and perform comparisons, such as the nonlocal structure comparison based method (Han et al., 2024b; Zhu et al., 2024), combining local and non-local structure based method (Chen et al., 2022, 2023b), adaptive graph based method (Han et al., 2024c), iterative sample augmentation network (Lv et al., 2024), dual-branch training method (Xing et al., 2024), self-guided autoencoders (Shi et al., 2024), commonality autoencoder (Wu et al., 2021), multiscale and enhanced UNet (Lv et al., 2022a, 2023), and self-supervised learning base method (Chen and Bruzzone, 2022). (iii) The image-translation based methods employ traditional regression functions or generative adversarial networks (GAN) to translate an image from one domain to the other, subsequently, compare the translated and original images within the target image domain to find the changes (Li et al., 2021; Jia et al., 2022a), such as image translation by using affinity matrix distance (Luppino et al., 2019), structure consistency based regression (Sun et al., 2023, 2024b), homogeneous pixel transformation (Liu et al., 2018b), fractal project base method (Mignotte, 2020), affinity matrix based X-Net and ACE-Net (Luppino et al., 2022), code-aligned autoencoder (Luppino et al., 2024), robust fusion GAN based method (Wang et al., 2024), multi-domain constrained translation network (Wu et al., 2024a), Copula mixtures guided translation network (Li et al., 2024), disentangled representation network (Dai et al., 2024), and graph convolutional network based HCD methods (Florez-Ospina et al., 2023; Han et al., 2024a).

Therefore, it can be found that for such comparison-based methods with $DI = \mathcal{M}_1(X) \ominus \mathcal{M}_2(Y)$, there are two obstacles to be crossed.

- **Firstly, how to ensure the accuracy of transformation functions of \mathcal{M}_1 and \mathcal{M}_2 ?** On one hand, \mathcal{M}_1 and \mathcal{M}_2 will be affected by the unknown changed regions (Sun et al., 2021a; Chen et al., 2023b), that is, it is inherently impossible to establish a one-to-one correspondence between the multitemporal images due to the intrinsic properties of change detection problems. Further,

the change influence consists of two aspects: first, during the construction or learning of \mathcal{M}_1 and \mathcal{M}_2 , and second, during the completion of the transformations of $\mathcal{M}_1(X)$ and $\mathcal{M}_2(Y)$ when the transformations are not based on individual pixels, i.e., the changed pixel will affect the surrounding unchanged pixels. Therefore, for unsupervised HCD methods, it is difficult to eradicate the negative impacts of unknown changing samples in the image transformation. For supervised HCD methods, although it is possible to establish accurate mapping relationships using labeled unchanged samples, this is achieved at the expense of a large amount of manual annotation. As is well known, accurately annotating changed regions on heterogeneous images requires strong expert knowledge (Lv et al., 2022b; Zheng et al., 2021).

- **Secondly, how to ensure robustness and generalizability of transformation functions of \mathcal{M}_1 and \mathcal{M}_2 ?** The comparison-based methods often rely on the expected imaging-invariant assumptions or need a large number of training samples to establish the correlation between heterogeneous images (Sun et al., 2022). On one hand, these assumed imaging invariances that hold for some simple scenarios may no longer be valid when faced with very complex HCD scenarios, such as when the image resolution is improved, the number of object classes is increased, the object relationships are complex, the image noise is exacerbated, etc. On the other hand, these transformation functions learned based on training samples also face the problems in carving out connections between heterogeneous images when the training samples are insufficient or contain erroneous samples. At the same time, they also face the generalization challenge, i.e., it is often difficult to directly apply the \mathcal{M}_1 and \mathcal{M}_2 obtained on one dataset to other datasets when imaging conditions such as imaging parameters and imaging scenes are different (Chen et al., 2024a). Therefore, these comparison-based HCD methods that rely on imaging-invariant assumptions or limited training samples struggle to transform complex heterogeneous images into a common domain for comparison, leading to a degradation in detection performance.

1.3. Motivation and contribution

Recently, Touati et al. have proposed some energy model based HCD methods (Touati and Mignotte, 2018; Touati et al., 2020). They have considered the pairwise relationships between each pixel and all other pixels, and established a nonlocal pairwise energy model (NLPEM) (Touati and Mignotte, 2018) to characterize the connections between pixel pair relationships and pixel pair labels. In their subsequent work (Touati et al., 2020), they have proposed another energy model (named as M3CD), which uses a metric function to calculate the probability that two pixels have different labels (one changed and the other unchanged) and those that have identical labels (both changed or unchanged). The attractiveness of M3CD lies in its use of pairwise constraints to construct an energy model for directly outputting the pixel label, i.e., the change map. Motivated by this pairwise relationship with pairwise labeling constraints, another energy model based on locality preservation (LPEM) is proposed (Sun et al., 2024a), which is devoted to constraining the multitemporal images to maintain the same local properties in the unchanged region, such as the local feature similarity and local spatial continuity. However, these methods still do not break away from the use of transformation and inter-image comparison in the metric function that measures the difference of pairs relationships.

Inspired by the Touati's energy model based methods (Touati and Mignotte, 2018; Touati et al., 2020) and the pairwise relationships based image structure used in previous methods (Sun et al., 2021b; Luppino et al., 2019), and to address the challenges of unsupervised HCD, we propose a simple yet highly effective approach, named rules induced energy model (RIEM for short). RIEM leverages the advantages

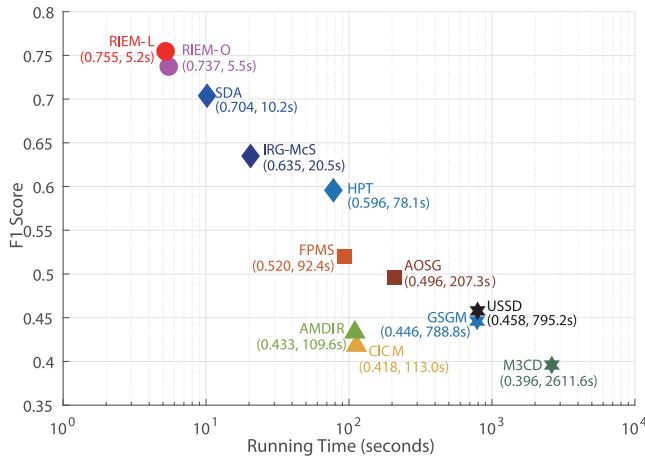


Fig. 1. Performance comparison between the proposed RIEM and representative existing methods, evaluated by average F1-score on seven datasets and typical runtime on images of size $2000 \times 2000 \times 3$.

of energy model and image structure simultaneously, which extends the concept that previous methods detect changes by measuring structure difference with graph mapping (Sun et al., 2021b) or direct comparison (Luppino et al., 2019). RIEM incorporates the structural associations between heterogeneous images as constraints into an energy model and extracts change information by solving the energy model. Most importantly, the six rules in RIEM used to determine changes are established based on the nature of the HCD problem without using any assumptions, and these rules provide a solution for detecting change without comparing multitemporal images, *i.e.*, the inter-image comparison of $\mathcal{M}_1(X) \ominus \mathcal{M}_2(Y)$ as most previous methods. Therefore, RIEM can address the challenges mentioned above and lead to a new solution for HCD problem (see Fig. 1).

Specifically, in RIEM, we first divide the multitemporal images into superpixels and choose each superpixel as a node. Then, we establish the connections between the class relationships (same/different) and change labels (changed/unchanged) of pairwise superpixels. And, we derive six rules for determining the change label of each superpixel from these connections, which can be further quantified as a function between the similarity/dissimilarity relationship and the change score (or probability) of pairwise superpixels, whereas the similarity and dissimilarity relationships between pairwise superpixels are captured by a k-nearest neighbor (KNN) graph and k-farthest neighbor (KFN) graph, respectively. Based on this, a rules induced energy model is built to characterize the connections between the change score of each superpixel and the structural and spatial correlations among pairwise superpixels, which implements four types of energy loss functions: a similarity relationships based loss, a dissimilarity relationships base loss, a spatial smoothness based loss, and a prior sparsity based loss. Finally, by solving this energy model, the change score and change label of each superpixel is obtained, then we can obtain both a difference image (DI) indicating the change score and a binary change map (CM) indicating the changed/unchanged areas.

Since the rules and constraints in the energy model is intuitively explainable and universal, the proposed RIEM is insensitive to imaging conditions, making it suitable for various HCD scenarios. Moreover, by using the rules and fully utilizing the flexibility of energy based model, RIEM can directly constrain the superpixel change labels, thereby obtaining the change results directly, without the processes of first transforming the images to the same domain and then comparing them as in the comparison-based methods. In addition, compared to the previous image structure base HCD methods that rely solely on similarity relationships, RIEM constructs a KFN graph to represent the dissimilarity relationships, complementing the similarity relationship

based KNN graph, which can more comprehensively capture the image structure. The main contributions of this work are as follows.

- We establish six rules for determining changes based on the nature of change detection problem without using any assumptions, which are intuitively explainable and universal across different scenarios. These rules enable detecting changes by considering only the intra-image relationships within each image without inter-image comparisons, liberating HCD from the complicated and challenging transformations and interactions between heterogeneous images.
- We convert the HCD to an energy minimization problem to release the ability of change determination rules, which characterizes the connections between change score (or change label) of each superpixel and relationships of pairwise superpixels.
- We fully exploit the relationship between superpixel pairs, including similarity relationships, dissimilarity relationships, and spatial neighborhood relationships, which enables more accurate constraints within the energy model.
- We conduct comparative experiments with the state-of-the-art methods on seven real datasets, and validate the accuracy and efficiency of the proposed HCD method. The code is released at <https://github.com/yulisun/RIEM>.

2. Connections and rules

Given two co-registered images collected at times t_1 (pre-change) and t_2 (post-change) by different sensors with different imaging conditions, denoted as $\mathbf{X} \in \mathbb{R}^{M \times N \times B_x}$ and $\mathbf{Y} \in \mathbb{R}^{M \times N \times B_y}$ respectively, whose pixels are denoted as $x(m, n, b_x)$ and $y(m, n, b_y)$ respectively. Here, $M \times N$ represents the image spatial size, and B_x (B_y) denotes the number of image channels. The goal of HCD is to obtain a binary change map that indicates whether each pixel has changed or not.

2.1. Pre-processing

As aforementioned, we have to consider the pairwise relationship within the image. Different from previous methods that choose pixel as basic unit (Touati and Mignotte, 2018), resulting a quadratic computational complexity of the number of pixels, *i.e.*, $\mathcal{O}(M^2N^2)$, we choose the superpixel as basic unit, which can effectively reduce complexity while utilizing contextual information and preserving object boundaries.

We employ the preprocessing method used in Sun et al. (2023) to obtain the co-segmented superpixels, which first concatenates the normalized (linear and logarithmic normalization for optical and SAR images respectively (Deledalle et al., 2012)) multitemporal images along the channel dimension, then uses the Gaussian mixture model based superpixel segmentation method (GMMS) (Ban et al., 2018) to segment the concatenated image into N_S superpixels, resulting in a segmentation map Λ . Using this Λ , we further segment the pre-change and post-change images into N_S superpixels \mathbf{X}_i and \mathbf{Y}_i , $i \in \mathcal{I}$ respectively, with the same edge contours, which are defined as

$$\begin{aligned}\mathbf{X}_i &= \{x(m, n, b_x) | (m, n) \in \Lambda_i, b_x = 1, \dots, B_x\} \\ \mathbf{Y}_i &= \{y(m, n, b_y) | (m, n) \in \Lambda_i, b_y = 1, \dots, B_y\}\end{aligned}\quad (1)$$

where index set is $\mathcal{I} = \{1, \dots, N_S\}$. In this case, the i th superpixels \mathbf{X}_i and \mathbf{Y}_i stand for the same geographical regions, and they are both internally homogeneous within themselves (representing the same kind of objects), respectively. So the pixels in the \mathbf{X}_i and \mathbf{Y}_i share the same change label, *i.e.*, all unchanged or changed.

2.2. Connections between pairwise relationship and change labels

We use O_i^{t1} and O_i^{t2} to denote the kinds of ground objects represented by the i th superpixels \mathbf{X}_i and \mathbf{Y}_i respectively, where $O_i^{t1} = O_i^{t2}$ and $O_i^{t1} \neq O_i^{t2}$ means that \mathbf{X}_i and \mathbf{Y}_i represent the same and different kinds of objects, respectively. We use L_i to denote the change label of the i th superpixel (\mathbf{X}_i and \mathbf{Y}_i) located at region A_i , where $L_i = 0$ represent the unchanged case and $L_i = 1$ represents the changed case. Then, we can divide the index set \mathcal{I} into changed subset $\mathcal{C} = \{i | L_i = 1\}$ and unchanged subset $\mathcal{U} = \{i | L_i = 0\}$. Next, we investigate the connection between the relationship and change labels of the pairwise superpixels.

For the pairwise superpixels of \mathbf{X}_i and \mathbf{X}_j , \mathbf{Y}_i and \mathbf{Y}_j , $i, j \in \mathcal{I}$, there are 4 combinations of relationships between pairwise superpixels within the same image expressed as

- $O_i^{t1} = O_j^{t1}, O_i^{t2} = O_j^{t2};$
 - $O_i^{t1} = O_j^{t1}, O_i^{t2} \neq O_j^{t2};$
 - $O_i^{t1} \neq O_j^{t1}, O_i^{t2} = O_j^{t2};$
 - $O_i^{t1} \neq O_j^{t1}, O_i^{t2} \neq O_j^{t2}.$
- (2)

For the pairwise superpixels of \mathbf{X}_i and \mathbf{Y}_i , \mathbf{X}_j and \mathbf{Y}_j , $i, j \in \mathcal{I}$, there are also 4 combinations of relationships between pairwise superpixels across different images, which also respectively equal to the pairwise change labels based on the nature of change detection problem:

- $O_i^{t1} = O_i^{t2}, O_j^{t1} = O_j^{t2} \Leftrightarrow L_i = 0, L_j = 0;$
 - $O_i^{t1} = O_i^{t2}, O_j^{t1} \neq O_j^{t2} \Leftrightarrow L_i = 0, L_j = 1;$
 - $O_i^{t1} \neq O_i^{t2}, O_j^{t1} = O_j^{t2} \Leftrightarrow L_i = 1, L_j = 0;$
 - $O_i^{t1} \neq O_i^{t2}, O_j^{t1} \neq O_j^{t2} \Leftrightarrow L_i = 1, L_j = 1.$
- (3)

Then, we can build the connection between the intra-image pairwise relationships and pairwise labels of the superpixels as follows:

$$\begin{cases} O_i^{t1} = O_j^{t1}; \\ O_i^{t2} = O_j^{t2}; \end{cases} \Rightarrow (L_i, L_j) \in \{(0, 0), (1, 1)\}; \quad (4a)$$

$$\begin{cases} O_i^{t1} = O_j^{t1}; \\ O_i^{t2} \neq O_j^{t2}; \end{cases} \Rightarrow (L_i, L_j) \in \{(0, 1), (1, 0), (1, 1)\}; \quad (4b)$$

$$\begin{cases} O_i^{t1} \neq O_j^{t1}; \\ O_i^{t2} = O_j^{t2}; \end{cases} \Rightarrow (L_i, L_j) \in \{(0, 1), (1, 0), (1, 1)\}; \quad (4c)$$

$$\begin{cases} O_i^{t1} \neq O_j^{t1}; \\ O_i^{t2} \neq O_j^{t2}; \end{cases} \Rightarrow (L_i, L_j) \in \{(0, 0), (0, 1), (1, 0), (1, 1)\}. \quad (4d)$$

In the (4a), when \mathbf{X}_i and \mathbf{X}_j represent the same kind of object (e.g. Forest), and \mathbf{Y}_i and \mathbf{Y}_j also represent the same kind of object (e.g. Forest or Water), then the labels of i th and j th superpixels can only be both unchanged ($L_i = 0, L_j = 0$) or changed ($L_i = 1, L_j = 1$). In the (4b) and (4c), when the superpixel pair represent the same kind of objects within one image and represent different kinds of objects within the other image, then the labels of i th and j th superpixels can never be both unchanged ($L_i = 0, L_j = 0$).

Although the connection between the pairwise relationships and pairwise labels has been considered in previous approaches, they are more or less incomplete and not as systematic and concise as (4a)–(4d) in this paper. For example, the structure consistency based methods (Sun et al., 2021b; Chen et al., 2022, 2023b; Florez-Ospina et al., 2023) are based on the fact that similarity relationships in unchanged regions are consistent, which only consider the connection of (4a) with “if $O_i^{t1} = O_j^{t1}$ and $L_i = L_j = 0$, then $O_i^{t2} = O_j^{t2}$ ”. The pairwise relationships based energy models (Touati and Mignotte, 2018; Touati et al., 2020) rely on the constraints expressed as: “ $O_i^{t1} = O_j^{t1}, O_i^{t2} = O_j^{t2} \Rightarrow L_i = L_j$ ” in the connection of (4a); and “other three pairwise relationships in (4b)–(4d) all lead to $L_i \neq L_j$ ”, which ignores some possible connections and thus affects the accuracy of energy model.

2.3. From connections to change determination rules

We established the corresponding connections between the relationships and labels of pairwise superpixels in (4a)–(4d), but none of them is a one-to-one mapping connection, i.e., it cannot be used directly to determine the change state of superpixels. However, when we impose certain restrictions on these relationships, we can derive rules for determining the change state, which can be used for the HCD problem.

Firstly, by comparing the (4a) and (4b), if we restrict \mathbf{X}_i and \mathbf{X}_j represent the same kind of objects, we can find that: (1) only if \mathbf{Y}_i and \mathbf{Y}_j also represent the same kind of object, it is possible to have the both unchanged $i, j \in \mathcal{U}$, that is, if \mathbf{Y}_i and \mathbf{Y}_j represent different kinds of objects, then at least one of the i th and j th superpixels changes; (2) only if \mathbf{Y}_i and \mathbf{Y}_j represent different kinds of objects, it is possible to have different labels for the i th and j th superpixel. Base on these, we can obtain the following two Rules.

RULE 1: when restricting $O_i^{t1} = O_j^{t1}$, if $O_i^{t2} \neq O_j^{t2}$ holds, then $(L_i, L_j) = (0, 0)$ does not hold.

RULE 2: when restricting $O_i^{t1} = O_j^{t1}$, if $O_i^{t2} = O_j^{t2}$ holds, then $L_i \neq L_j$ does not hold.

Secondly, by comparing (4a) and (4c), we can similarly obtain two Rules as follows.

RULE 3: when restricting $O_i^{t2} = O_j^{t2}$, if $O_i^{t1} \neq O_j^{t1}$ holds, then $(L_i, L_j) = (0, 0)$ does not hold.

RULE 4: when restricting $O_i^{t2} = O_j^{t2}$, if $O_i^{t1} = O_j^{t1}$ holds, then $L_i \neq L_j$ does not hold.

Thirdly, by comparing (4c) and (4d), if we restrict \mathbf{X}_i and \mathbf{X}_j represent different kinds of objects, we can find that only if \mathbf{Y}_i and \mathbf{Y}_j also represents different kinds of objects, it is possible to have both unchanged $i, j \in \mathcal{U}$, that is, if \mathbf{Y}_i and \mathbf{Y}_j represent the same kind of objects, then at least one of the i th and j th superpixels changes. Base on this, we can obtain the Rule 5 as

RULE 5: when restricting $O_i^{t1} \neq O_j^{t1}$, if $O_i^{t2} = O_j^{t2}$ holds, then $(L_i, L_j) = (0, 0)$ does not hold.

Fourth, by comparing (4b) and (4d), we can similarly obtain another Rule 6 as

RULE 6: when restricting $O_i^{t2} \neq O_j^{t2}$, if $O_i^{t1} = O_j^{t1}$ holds, then $(L_i, L_j) = (0, 0)$ does not hold.

In Table 1, we use simple symbolic examples to visually illustrate the connections between pairwise relationships and change labels, and demonstrate how these connections are used to derive the rules for determining the change state (only Rules 1, 2, and 5 are shown to avoid repetition). There are two particularly appealing aspects of these six Rules.

First, these Rules are derived based on the nature of change detection problem and no other assumptions are used, so they are universal across different scenarios, for both homogeneous and heterogeneous change detection problems.

Table 1

The process of Rule derivation from the connections between pairwise relationships and change labels.

Rules	Restrictions	Criteria	Examples $(\mathbf{X}_i, \mathbf{X}_j) \rightarrow (\mathbf{Y}_i, \mathbf{Y}_j)$	Lables	Conclusions
Rule 1	$O_i^{t1} = O_j^{t1}$	$O_i^{t2} \neq O_j^{t2}$	$\bullet\bullet \rightarrow \bullet\Delta$	$(L_i, L_j) = (0, 1)$	$(L_i, L_j) = (0, 0)$ does not hold
			$\bullet\bullet \rightarrow \Delta\bullet$	$(L_i, L_j) = (1, 0)$	
			$\bullet\bullet \rightarrow \Delta\star$	$(L_i, L_j) = (1, 1)$	
Rule 2	$O_i^{t1} = O_j^{t1}$	$O_i^{t2} = O_j^{t2}$	$\bullet\bullet \rightarrow \bullet\bullet$	$(L_i, L_j) = (0, 0)$	$L_i \neq L_j$ does not hold
			$\bullet\bullet \rightarrow \Delta\Delta$	$(L_i, L_j) = (1, 1)$	
			$\bullet\Delta \rightarrow \bullet\bullet$	$(L_i, L_j) = (0, 1)$	
Rule 5	$O_i^{t1} \neq O_j^{t1}$	$O_i^{t2} = O_j^{t2}$	$\bullet\Delta \rightarrow \Delta\Delta$	$(L_i, L_j) = (1, 0)$	$(L_i, L_j) = (0, 0)$ does not hold
			$\bullet\Delta \rightarrow \star\star$	$(L_i, L_j) = (1, 1)$	

The symbols $\bullet\Delta\star$ represent different kinds of objects, e.g., \bullet Forest, Δ Water, \star Building.

- Second, these Rules provide a nice solution to the HCD problem, enabling to detect changes by considering only the intra-image relationships within each image, without considering inter-image comparisons, which greatly reduces the difficulty of solving the HCD problem, since the former avoids the transformations and interactions between heterogeneous images.

Nevertheless, how to utilize these Rules to detect changes is to be addressed, that is, how to impose the restrictions of these Rules and how to mathematically implement these inferences for the HCD problem. Next, we utilize the energy model to address these challenges, which can encode the Rules by building different energy functions.

3. Rules induced energy model

Energy based model (EBM) is a mathematical framework that assigns a scalar energy value to each possible configuration of variables for capturing their dependencies (LeCun et al., 2006), where lower energy corresponds to correct or desirable configurations and higher energy corresponds to incorrect or undesirable configurations. The advantage of using an EBM lies in its flexibility and expressiveness, as the energy can be constructed in a logical manner according to specific requirements (Isack and Boykov, 2012; Agoritas et al., 2023; Delong et al., 2012). This characteristic has also been leveraged in the change detection problem (Touati and Mignotte, 2018; Touati et al., 2020). In this study, we employ an EBM to release the potential benefits of the Rules 1–6, as illustrated by Fig. 2.

We define the change vector to be solved as $\mathbf{p} \in \mathbb{R}^{N_S}$, where $0 \leq p_i \leq 1$ denotes the change score of the i th superpixel, with $p_i = 0$ representing “unchanged” and $p_i = 1$ representing “changed”. Then we can obtain the energy model as

$$\mathbf{p}^* = \arg \min_{0 \leq p_i \leq 1} E_h(\mathbf{p}; \mathbf{X}, \mathbf{Y}) \quad (5)$$

where the energy function $E_h(\mathbf{p}; \mathbf{X}, \mathbf{Y})$ can be used to measure the plausibility of different configurations of values for \mathbf{X} , \mathbf{Y} , and \mathbf{p} . It assigns a scalar energy value to each configuration, indicating how well it aligns with the desired constraints or objectives of the problem. By minimizing the energy function, we can identify the most likely configuration of \mathbf{X} , \mathbf{Y} , and \mathbf{p} that best satisfies the given constraints and achieves the desired outcome in the context of the problem at hand.

3.1. Similarity relationship base energy

3.1.1. KNN graph

In the Rules 1–2, their restrictions are $O_i^{t1} = O_j^{t1}$, i.e., \mathbf{X}_i and \mathbf{X}_j refer to the same kind of objects. To meet this restriction, we construct

KNN graph $G_n^{t1} = \{V_n^{t1}, E_n^{t1}, A_n^{t1}\}$ for the pre-change image to capture the similarity relationship within the image \mathbf{X} as

$$\begin{aligned} V_n^{t1} &= \{i | i \in \mathcal{I}\}; \\ E_n^{t1} &= \{(i, j) | j \in \mathcal{N}_i^x, i \in \mathcal{I}\} \end{aligned} \quad (6)$$

where V_n^{t1} is the vertex set, E_n^{t1} is the edge set, A_n^{t1} denotes the adjacent matrix, \mathcal{N}_i^x denotes the index of the k nearest-neighbor (NN) of \mathbf{X}_i . Define \mathbf{D}^x as the superpixel feature distance matrix of pre-change image, whose element $D_{i,j}^x$ denotes the feature distance between the i th and j th superpixels. Then, \mathcal{N}_i^x can be constructed by sorting the distance vector \mathbf{D}_i^x and taking out the index of first k minima other than i . Then, based on the properties of the KNN graph (when an appropriate k is chosen), it can be assumed that the connected vertices (*i.e.*, superpixels) on the graph represent the same kind of objects. Similarly, we can also construct a KNN graph $G_n^{t2} = \{V_n^{t2}, E_n^{t2}, A_n^{t2}\}$ for the post-change image by using the superpixel feature distance matrix \mathbf{D}^y .

The choice of k in the KNN graph is important. On one hand, k should not be too large to avoid connecting superpixels representing different kinds of objects, which would lead to the case that there exists an $(i, j) \in E_n^{t1}$ such that $O_i^{t1} \neq O_j^{t1}$. On the other hand, k should not be too small, as a small number of edges in the KNN graph (*i.e.*, $|E_n^{t1}|$ is small) would not sufficiently capture the similarity structure of the image, resulting in limited contribution for Rule 1.

We first use a small k as $k = \sqrt{N_S}$ (Loftsgaarden and Quesenberry, 1965; Fukunaga and Hostetler, 1973; Mitra et al., 2002) to construct the initial KNN graph, and then employ a neighbor expansion strategy based on the well-known balance theory in the social psychology (Heider, 1946; Cartwright and Harary, 1956): which posits that social relationships adhere four rules: “the friend of my friend is my friend”, “the enemy of my friend is my enemy”, “the friend of my enemy is my enemy” and “the enemy of my enemy is my friend”. With the first rule in balance theory, we employ “the nearest-neighbors of nearest-neighbors are also considered as nearest-neighbors” to obtain a high-order KNN graph $G_{hn}^{t1} = \{V_{hn}^{t1}, E_{hn}^{t1}, A_{hn}^{t1}\}$ connecting more pairwise superpixels truly representing the same kind of objects. For example, the edge set E_{2n}^{t1} of the second-order graph is defined as

$$(i, j) \in E_{2n}^{t1}; \quad \text{if } (i, j) \in E_n^{t1} \quad (7) \\ \text{or } (i, t) \in E_n^{t1}, (t, j) \in E_n^{t1}$$

and the edge set E_{3n}^{t1} of the third-order graph is

$$(i, j) \in E_{3n}^{t1}; \quad \text{if } (i, j) \in E_{2n}^{t1} \quad (8) \\ \text{or } (i, t) \in E_{2n}^{t1}, (t, j) \in E_n^{t1}$$

Similarly, we can also construct a high-order KNN graph $G_{hn}^{t2} = \{V_{hn}^{t2}, E_{hn}^{t2}, A_{hn}^{t2}\}$ for the post-change image. In the proposed method, we utilize the third-order KNN graphs.

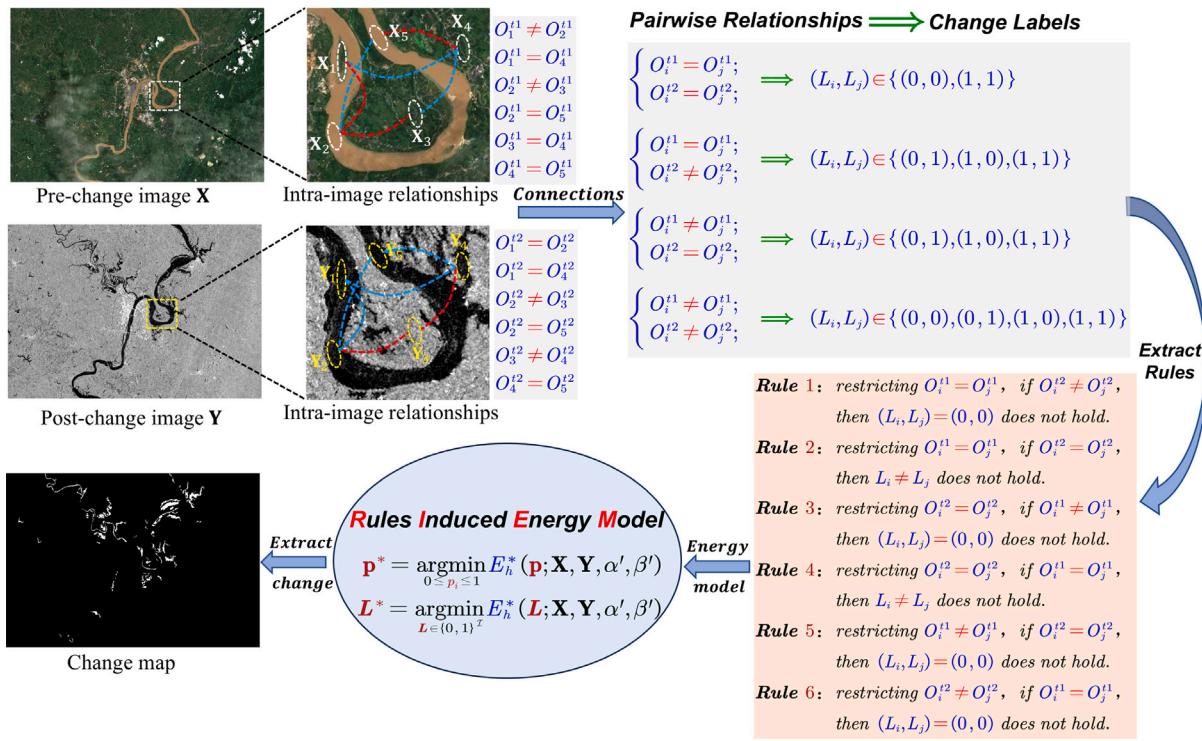


Fig. 2. Illustration of the proposed rules-induced energy model.

3.1.2. Energy functions

For any adjacent vertices on G_{hn}^{t1} , i.e., $(i, j) \in E_{hn}^{t1}$, by using Rule 1, the greater the probability that $O_i^{t2} \neq O_j^{t2}$ holds, the less likely $L_i = 0, L_j = 0$ is to occur, then we can build the following energy function for Rule 1

$$E_1 = \sum_{(i,j) \in E_{hn}^{t1}} (1 - p_i) (1 - p_j) f_{i,j}^1 \quad (9)$$

with

$$f_{i,j}^1 = \begin{cases} D_{i,j}^y, & \text{if } (i, j) \notin E_{hn}^{t2} \\ 0, & \text{if } (i, j) \in E_{hn}^{t2} \end{cases} \quad (10)$$

where $f_{i,j}^1$ can be used to measure the energy when the i th and j th vertices are both unchanged ($p_i = p_j = 0$). As can be seen from (9) and (10), if the feature difference between \mathbf{Y}_i and \mathbf{Y}_j is large (i.e., a big $D_{i,j}^y$), which means that \mathbf{Y}_i and \mathbf{Y}_j most likely belong to different kinds of objects, then to decrease the value of E_1 , larger values of p_i or p_j are required, indicating changed labels for i or j with $(i, j) \in E_{hn}^{t1}$.

For any adjacent vertices on G_{hn}^{t1} , i.e., $(i, j) \in E_{hn}^{t1}$, by using Rule 2, the greater the probability that $O_i^{t2} = O_j^{t2}$ holds, the less likely $L_i \neq L_j$ is to occur, then we can build the following energy function for Rule 2

$$E_2 = \sum_{(i,j) \in E_{hn}^{t1}} (p_i - p_j)^2 f_{i,j}^2 \quad (11)$$

with

$$f_{i,j}^2 = \begin{cases} \exp(-D_{i,j}^y), & \text{if } (i, j) \in E_{hn}^{t2} \\ 0, & \text{if } (i, j) \notin E_{hn}^{t2} \end{cases} \quad (12)$$

where $f_{i,j}^2$ can be used to measure the energy when the i th and j th vertices have different change labels ($p_i \neq p_j$). As can be seen from (11) and (12), if the feature difference between \mathbf{Y}_i and \mathbf{Y}_j is small (i.e., a big $f_{i,j}^2$), which means that \mathbf{Y}_i and \mathbf{Y}_j most likely refer to the same kind of object, then to decrease the value of E_2 , smaller value of $(p_i - p_j)^2$ is required, indicating the same change label of i and j with $(i, j) \in E_{hn}^{t1}$.

Similarly, in order to utilize the Rule 3 and Rule 4, we consider the adjacent vertices on G_{hn}^{t2} , i.e., $(i, j) \in E_{hn}^{t2}$, and build the following energy functions of E_3 and E_4 for Rule 3 and Rule 4, respectively.

$$\begin{aligned} E_3 &= \sum_{(i,j) \in E_{hn}^{t2}} (1 - p_i) (1 - p_j) f_{i,j}^3 \\ E_4 &= \sum_{(i,j) \in E_{hn}^{t2}} (p_i - p_j)^2 f_{i,j}^4 \end{aligned} \quad (13)$$

with

$$\begin{aligned} f_{i,j}^3 &= \begin{cases} D_{i,j}^x, & \text{if } (i, j) \notin E_{hn}^{t1} \\ 0, & \text{if } (i, j) \in E_{hn}^{t1} \end{cases} \\ f_{i,j}^4 &= \begin{cases} \exp(-D_{i,j}^x), & \text{if } (i, j) \in E_{hn}^{t1} \\ 0, & \text{if } (i, j) \notin E_{hn}^{t1} \end{cases} \end{aligned} \quad (14)$$

where $(i, j) \in E_{hn}^{t2}$ in (13) is used to satisfy the restriction of $O_i^{t2} = O_j^{t2}$, $f_{i,j}^3$ is used to measure the energy when the i th and j th vertices are both unchanged ($p_i = p_j = 0$), $f_{i,j}^4$ is used to measure the energy when the i th and j th vertices have different change labels ($p_i \neq p_j$).

From the Rules 1 to 4 and the energy functions of E_1 to E_4 , we can find that they utilize the similarity relationships in the image, specifically, they require that the similarity structure in one image represented by the KNN graph in unchanged regions should be preserved by the other image (e.g. E_1 and E_3). By combining these energy functions of E_1 to E_4 , we can obtain the similarity relationships based energy function E_{sr} as

$$E_{sr} = E_1 + E_2 + E_3 + E_4 \quad (15)$$

3.2. Dissimilarity relationship base energy

3.2.1. KFN graph

In the Rule 5, its restriction is $O_i^{t1} \neq O_j^{t1}$, i.e., \mathbf{X}_i and \mathbf{X}_j refer to different kinds of objects. To meet this restriction, we construct KFN graph of $\mathcal{G}_f^{t1} = \{\mathcal{V}_f^{t1}, \mathcal{E}_f^{t1}, \mathcal{A}_f^{t1}\}$ for pre-change image to capture the

dissimilarity relationships within \mathbf{X} as

$$\begin{aligned} \mathcal{V}_f^{t1} &= \{i | i \in \mathcal{I}\}; \\ \mathcal{E}_f^{t1} &= \{(i, j) | j \in \mathcal{F}_i^x, i \in \mathcal{I}\} \end{aligned} \quad (16)$$

where \mathcal{F}_i^x denotes the index of k farthest-neighbor (FN) of \mathbf{X}_i , i.e., $(i, j) \in \mathcal{E}_f^{t1}$ if and only if $D_{i,j}^x$ is one of the k -maxima of feature distance vector \mathbf{D}_i^x . Then, based on the properties of the KFN graph (when an appropriate k is chosen), it can be assumed that connected vertices (i.e., superpixels) on the graph represent the different kinds of objects. Similarly, we can also construct a KFN graph $\mathcal{G}_f^{t2} = \{\mathcal{V}_f^{t2}, \mathcal{E}_f^{t2}, \mathcal{A}_f^{t2}\}$ for the post-change image \mathbf{Y} .

Different from the KNN graph that captures similarity relationships, constructing the KFN graph to represent dissimilarity relationships requires paying attention on two aspects.

- First, for the selection of k . It is well known that remote sensing images contain a wide range of land cover classes, so the number of superpixel pairs representing different kinds of objects in the image will be much larger than the number of superpixel pairs representing the same kind of objects. Therefore the number of edges in KFN graph should be larger than that in KNN graph. In the proposed RIEM, we select $k = 5\sqrt{N_S}$ for the initial KFN graph.
- Second, for the neighbors expansion. At this time, the strategy of “the neighbors of neighbors are also considered as neighbors” used in KNN graph is no longer suitable for KFN graph. For example, suppose that \mathbf{X}_i represents the object of “grass” and it connects \mathbf{X}_j and \mathbf{X}_l , which both represent the “water” in the initial KFN graph, by using the neighbor expansion strategy in KNN graph, at this point \mathbf{X}_j and \mathbf{X}_l that belong to the same kind of object will be connected by the high-order KFN graph. Here, we use another strategy for neighbor expansion of KFN graph: “**the nearest-neighbor of the farthest-neighbor and the farthest-neighbor of the nearest-neighbor are also likely to be the farthest-neighbor**”, which are similar to the second and third rules in balance theory (Heider, 1946; Cartwright and Harary, 1956).

Then we can construct the KFN graph $\mathcal{G}_{hf}^{t1} = \{\mathcal{V}_{hf}^{t1}, \mathcal{E}_{hf}^{t1}, \mathcal{A}_{hf}^{t1}\}$ as follows

$$\begin{aligned} (i, j) \in \mathcal{E}_{hf}^{t1}; \quad &\text{if } (i, j) \in \mathcal{E}_f^{t1} \\ &\text{or } (i, t) \in \mathcal{E}_f^{t1}, (t, j) \in \mathcal{E}_{hn}^{t1} \\ &\text{or } (i, t) \in \mathcal{E}_{hn}^{t1}, (t, j) \in \mathcal{E}_f^{t1} \end{aligned} \quad (17)$$

Similarly, we can also construct a high-order KFN graph $\mathcal{G}_{hf}^{t2} = \{\mathcal{V}_{hf}^{t2}, \mathcal{E}_{hf}^{t2}, \mathcal{A}_{hf}^{t2}\}$ for the post-change image. With the help of this neighbors expansion strategy, the constructed high-order KFN graphs can connect more truly pairwise superpixels representing different kinds of objects. Fig. 3 illustrates the neighbor expansion strategy for high-order graphs.

3.2.2. Energy functions

For any adjacent vertices on \mathcal{G}_{hf}^{t1} , i.e., $(i, j) \in \mathcal{E}_{hf}^{t1}$, by using Rule 5, the greater the probability that $O_i^{t2} = O_j^{t2}$ holds, the less likely $L_i = 0$, $L_j = 0$ is to occur, then we can obtain the following energy function for Rule 5

$$E_5 = \sum_{(i,j) \in \mathcal{E}_{hf}^{t1}} (1 - p_i)(1 - p_j) f_{i,j}^2 \quad (18)$$

where $f_{i,j}^2$ is defined in (12) and can be used to measure the energy when the i th and j th vertices are both unchanged ($p_i = p_j = 0$) in (18). As can be seen from (18), if the \mathbf{Y}_i and \mathbf{Y}_j are quite similar (i.e., a big $f_{i,j}^2$), that is, \mathbf{Y}_i and \mathbf{Y}_j are likely to belong to the same kind of objects,

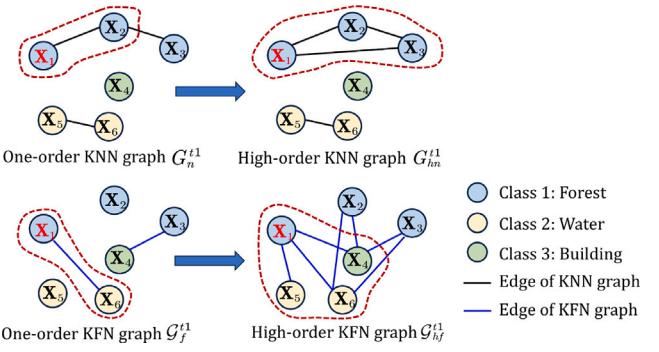


Fig. 3. Illustration of the neighbor expansion strategy for high-order graphs.

then to decrease the energy loss E_5 , smaller value of $(1 - p_i)(1 - p_j)$ is required, indicating changed labels for i or j with $(i, j) \in \mathcal{E}_{hf}^{t1}$.

Similarly, for the Rule 6, we consider the adjacent vertices on \mathcal{G}_{hf}^{t2} , i.e., $(i, j) \in \mathcal{E}_{hf}^{t2}$, we can build the following energy function

$$E_6 = \sum_{(i,j) \in \mathcal{E}_{hf}^{t2}} (1 - p_i)(1 - p_j) f_{i,j}^4 \quad (19)$$

where $(i, j) \in \mathcal{E}_{hf}^{t2}$ is used to satisfy the restriction of $O_i^{t2} \neq O_j^{t2}$ in Rule 6, and $f_{i,j}^4$ is defined in (14) to measure the energy when the i th and j th vertices are both unchanged ($p_i = p_j = 0$) in (19).

From the Rules 5–6 and the energy functions of E_5 and E_6 , we can find that they utilize the dissimilarity relationships in the image, specifically, they require that the dissimilarity structure in one image represented by the KFN graph in unchanged regions should be preserved by the other image. By combining the energy functions of E_5 and E_6 , we can obtain the dissimilarity relationships based energy function E_{dr} as

$$E_{dr} = E_5 + E_6 \quad (20)$$

3.3. Spatial smoothness based energy

According to the first law of geography, spatially adjacent objects are similar and correlated to a certain extent. Therefore, it is reasonable to assume that the change states of spatially closed superpixels are correlated. This means that if one superpixel undergoes a change, its spatially neighboring superpixel are more likely to undergo changes as well. This hypothesis is widely applied in change detection to infer and predict the propagation and diffusion of changes, which reduces the salt-and-pepper noise in change map. In this paper, we use a spatial smoothness based energy similar as the discontinuity penalty used in Sun et al. (2021a).

Firstly, we construct spatial neighbors for each superpixel by considering two conditions: if the i th and j th superpixel edges are intersecting or if the distance between their centroids is smaller than a predefined threshold R , then we consider them to be spatially close neighbors, defined as $(i, j) \in \mathcal{N}^R$. As the average superpixel size is MN/N_S , the R in RIEM is simply set to $2\sqrt{MN/N_S}$.

Secondly, we construct an energy function E_{ss} based on spatial smoothness, which incorporates both traditional spatial context information as well as similarity/dissimilarity relationships in multitemporal images involved in the HCD problem.

$$E_{ss} = \sum_{(i,j) \in \mathcal{N}^R} \frac{\phi_{i,j}(p_i - p_j)^2}{d_{i,j}} \quad (21)$$

where $d_{i,j}$ denotes the spatial distance between two superpixels located at Λ_i and Λ_j (i.e. spatial distance between center pixels), and the

penalty function $\phi_{i,j}$ is defined as

$$\phi_{i,j} = \begin{cases} 1/2, & \text{if } D_{i,j}^x > \rho_1, D_{i,j}^y > \rho_2 \\ \sigma \left(\frac{2(D_{i,j}^x - \rho_1)(D_{i,j}^y - \rho_2)}{\rho_1 \rho_2} \right), & \text{otherwise} \end{cases} \quad (22)$$

Here $\rho_1 = \frac{\langle D_{i,j}^x \rangle_{(i,j) \in E_{hn}^{t1}} + \langle D_{i,j}^x \rangle_{(i,j) \in E_{hf}^{t1}}}{2}$, $\rho_2 = \frac{\langle D_{i,j}^y \rangle_{(i,j) \in E_{hn}^{t2}} + \langle D_{i,j}^y \rangle_{(i,j) \in E_{hf}^{t2}}}{2}$ represent the average neighborhood feature distance with $\langle \cdot \rangle$ denoting the average operation, and σ represents the sigmoid function of $\sigma(z) = 1/(1 + e^{-z})$. As can be seen from (21) and (22), the penalty function $\phi_{i,j}$ punishes the non-smoothness of $p_i \neq p_j$, $(i, j) \in \mathcal{N}^R$ with four different cases:

(1) When $D_{i,j}^x \leq \rho_1$ and $D_{i,j}^y \leq \rho_2$, we consider it tends to the situation of $O_i^{t1} = O_j^{t1}$ and $O_i^{t2} = O_j^{t2}$, based on the pairwise connection of (4a), $\phi_{i,j}$ gives a large penalty on the non-smoothness of $p_i \neq p_j$, and as $D_{i,j}^x$ and $D_{i,j}^y$ decrease, the non-smoothness penalty is larger.

(2) When $D_{i,j}^x \leq \rho_1$ and $D_{i,j}^y > \rho_2$, we consider it tends to the situation of $O_i^{t1} = O_j^{t1}$ and $O_i^{t2} \neq O_j^{t2}$. In this case, the probability that both the i th and j th superpixels are changing is relatively small. That is, it is rare for spatially adjacent superpixels ($(i, j) \in \mathcal{N}^R$) that belong to the same kind of objects (such as “grass”) to change into different kinds of objects (such as “buildings” and “water”, respectively) in a single event. Therefore, based on the pairwise connection of (4b), $\phi_{i,j}$ gives a small penalty on the non-smoothness of $p_i \neq p_j$, and as $D_{i,j}^x$ decrease and $D_{i,j}^y$ increase, the non-smoothness penalty is smaller.

(3) When $D_{i,j}^x > \rho_1$ and $D_{i,j}^y \leq \rho_2$, we consider it tends to the situation of $O_i^{t1} \neq O_j^{t1}$ and $O_i^{t2} = O_j^{t2}$. Similar to case 2, $\phi_{i,j}$ gives a small penalty on the non-smoothness of $p_i \neq p_j$ based on the pairwise connection of (4c).

(4) When $D_{i,j}^x > \rho_1$ and $D_{i,j}^y > \rho_2$, we consider it tends to the situation of $O_i^{t1} \neq O_j^{t1}$ and $O_i^{t2} \neq O_j^{t2}$. In this case, the spatially adjacent superpixels may be located at the boundary between different kinds of objects, which predicts that the two superpixels are not closely related. Then, based on the pairwise connection of (4d), $\phi_{i,j}$ gives a median penalty of $\phi_{i,j} = 1/2$.

3.4. Prior sparsity based energy

In practical change detection problems, we have a prior knowledge about the sparsity of changes. In real-world scenarios, only a small portion of regions undergo changes, while the majority of regions remain unchanged. This sparse prior knowledge is widely applied in other change detection methods (Sun et al., 2023; Touati and Mignotte, 2018; Touati et al., 2020). Theoretically, we can use the ℓ_0 -norm-based $\|\mathbf{p}\|_0$ as a sparse penalty function. However, considering the non-convexity of the ℓ_0 -norm, we use its convex relaxation of ℓ_1 -norm penalty. Therefore, we obtain an energy function E_{ps} based on the prior sparsity knowledge as

$$E_{ps} = \sum_{i \in I} p_i \quad (23)$$

which discourages changes.

3.5. Hybrid energy model

By incorporating the energies of E_{sp} (15), E_{dr} (20), E_{ss} (21) and E_{ps} (23), we convert the HCD problem to the following hybrid energy function E_h as (see the Eq. (24) which is given in Box I).

We define the matrices of $\mathbf{B}^{(1)}$, $\mathbf{B}^{(2)}$, $\mathbf{W}^{(1)}$ and $\mathbf{W}^{(2)}$, whose elements are

$$\begin{aligned} B_{i,j}^{(1)} &= f_{i,j}^1 \delta((i, j) \in E_{hn}^{t1}) + f_{i,j}^3 \delta((i, j) \in E_{hn}^{t2}) \\ B_{i,j}^{(2)} &= f_{i,j}^2 \delta((i, j) \in E_{hf}^{t1}) + f_{i,j}^4 \delta((i, j) \in E_{hf}^{t2}) \\ W_{i,j}^{(1)} &= f_{i,j}^2 \delta((i, j) \in E_{hn}^{t1}) + f_{i,j}^4 \delta((i, j) \in E_{hn}^{t2}) \\ W_{i,j}^{(2)} &= \frac{\phi_{i,j}}{d_{i,j}} \delta((i, j) \in \mathcal{N}^R) \end{aligned} \quad (25)$$

where $\delta(\cdot)$ is 1 when the condition in parentheses holds, and 0 otherwise. Substitute (25) into (24), we have

$$\begin{aligned} E_h = & \sum_{i,j \in I} (1 - p_i) (1 - p_j) (B_{i,j}^{(1)} + B_{i,j}^{(2)}) \\ & + \sum_{i,j \in I} (p_i - p_j)^2 (W_{i,j}^{(1)} + W_{i,j}^{(2)}) + \sum_{i \in I} p_i \end{aligned} \quad (26)$$

In order to maintain a balance between the various energy components within E_h , we define the matrices of $\mathbf{B} = \mathbf{B}^{(1)} + \frac{\sum_{i,j=1}^{N_S} B_{i,j}^{(1)}}{\sum_{i,j=1}^{N_S} B_{i,j}^{(2)}} \mathbf{B}^{(2)}$, $\mathbf{W} = \mathbf{W}^{(1)} + \frac{\sum_{i,j=1}^{N_S} W_{i,j}^{(1)}}{\sum_{i,j=1}^{N_S} W_{i,j}^{(2)}} \mathbf{W}^{(2)}$, and define the Laplacian matrix \mathbf{L}_w as $\mathbf{L}_w = \mathbf{D}_w - (\mathbf{W} + \mathbf{W}^T)/2$ with $\mathbf{D}_w \in \mathbb{R}^{N_S \times N_S}$ being a diagonal matrix whose i th diagonal element is $\sum_{j=1}^{N_S} (W_{i,j} + W_{j,i})/2$, then we introduce two weighting parameters and modify E_h to E_h^* ($\mathbf{p}; \mathbf{X}, \mathbf{Y}, \alpha, \beta$) as

$$\begin{aligned} E_h^*(\mathbf{p}; \mathbf{X}, \mathbf{Y}, \alpha, \beta) = & \sum_{i,j \in I} (1 - p_i) (1 - p_j) B_{i,j} \\ & + \alpha/2 \sum_{i,j \in I} (p_i - p_j)^2 W_{i,j} + \beta \sum_{i \in I} p_i \\ & = (\mathbf{1} - \mathbf{p})^T \mathbf{B} (\mathbf{1} - \mathbf{p}) + \alpha \mathbf{p}^T \mathbf{L}_w \mathbf{p} + \beta \mathbf{p}^T \mathbf{1} \end{aligned} \quad (27)$$

where $\mathbf{1}$ an N_S -dimensional column vector with all elements equal to 1. We further set the weighting parameters as

$$\alpha = \frac{\alpha' \sum_{i,j} B_{i,j}}{\sum_{i,j} W_{i,j}}; \quad \beta = \frac{\beta' \sum_{i,j} B_{i,j}}{N_S} \quad (28)$$

where α' and β' are easier to adjust. Then, we can obtain the final energy model (5) for the HCD problem as

$$\mathbf{p}^* = \arg \min_{0 \leq p_i \leq 1} E_h^*(\mathbf{p}; \mathbf{X}, \mathbf{Y}, \alpha', \beta') \quad (29)$$

In the E_h^* of (27), when $B_{i,j}$ is larger, it indicates that the inconsistency of similarity/dissimilarity relationships across multitemporal images is greater. To reduce the energy value E_h^* , it tends to obtain larger values of p_i or p_j , i.e., higher change scores for the i th or j th superpixels. Therefore, the meaning of the energy function $(\mathbf{1} - \mathbf{p})^T \mathbf{B} (\mathbf{1} - \mathbf{p})$ is that when the image structures represented by similarity/dissimilarity relationships are inconsistent, for example, if \mathbf{X}_i and \mathbf{X}_j are similar but \mathbf{Y}_i and \mathbf{Y}_j are dissimilar, or if \mathbf{X}_i and \mathbf{X}_j are dissimilar but \mathbf{Y}_i and \mathbf{Y}_j are similar, then the energy loss can be utilized to detect the changes. In the E_h^* , the energy function $\mathbf{p}^T \mathbf{L}_w \mathbf{p}$ indicates that the change vector \mathbf{p} is smooth on the graph G_s constructed as

$$\begin{aligned} V_s &= \{i | i \in I\}; \\ E_s &= \mathcal{N}^R \cup \{E_{hn}^{t1} \cap E_{hn}^{t2}\} \end{aligned} \quad (30)$$

From the E_h^* of (27), we can also observe that the energy functions are antagonistic to each other. The first term tends to obtain a changed \mathbf{p} , i.e., $\mathbf{p} = \mathbf{1}$, the second term tends to obtain a smooth \mathbf{p} , i.e., $\mathbf{p} = c\mathbf{1}$ with c being a constant, while the last term tends to obtain the unchanged \mathbf{p} , i.e., $\mathbf{p} = \mathbf{0}$. Additionally, we can see that E_h^* is directly built on the superpixel change states (i.e., solving for \mathbf{p} directly), which is simple in form and has clear physical meaning and interpretability.

3.6. Change extraction

Because of the simple form of the objective function E_h^* (27), the energy optimization problem (29) could be approximately solved by the gradient descent method. Since the value of $B_{i,j}$ can partly reflect the likelihood of occurring changes in the i th and j th superpixels according to (25) and (27), we set the initial change vector \mathbf{p}^0 to the normalized $\frac{\mathbf{B}^1 + \mathbf{B}^T \mathbf{1}}{\max(\mathbf{B}^1 + \mathbf{B}^T \mathbf{1})}$. Once the solution \mathbf{p}^* is obtained, we can gain a DI by using p_i^* to indicate the change score of each pixel within the i th superpixel as

$$\text{DI}(m, n) = p_i^*; \quad (m, n) \in \Lambda_i, \quad i = 1, \dots, N_S \quad (31)$$

$$\begin{aligned}
E_h = & E_{sr} + E_{dr} + E_{ss} + E_{ps} \\
= & \underbrace{\sum_{(i,j) \in E_{hn}^{t1}} (1-p_i)(1-p_j)f_{i,j}^1 + \sum_{(i,j) \in E_{hn}^{t1}} (p_i - p_j)^2 f_{i,j}^2 + \sum_{(i,j) \in E_{hn}^{t2}} (1-p_i)(1-p_j)f_{i,j}^3 + \sum_{(i,j) \in E_{hn}^{t2}} (p_i - p_j)^2 f_{i,j}^4}_{E_{sr}} \\
& + \underbrace{\sum_{(i,j) \in E_{hf}^{t1}} (1-p_i)(1-p_j)f_{i,j}^2 + \sum_{(i,j) \in E_{hf}^{t2}} (1-p_i)(1-p_j)f_{i,j}^4}_{E_{dr}} + \underbrace{\sum_{(i,j) \in \mathcal{N}^R} \frac{\phi_{i,j}(p_i - p_j)^2}{d_{i,j}}}_{E_{ss}} + \underbrace{\sum_{i \in I} p_i}_{E_{ps}}
\end{aligned} \tag{24}$$

Box I.

In order to obtain a binary CM, we have two strategies. First, we compute CM by dividing \mathbf{p}^* into “changed” and “unchanged” classes with various segmentation methods, such as thresholding (e.g., Otsu’s thresholding (Otsu, 1979)), clustering (e.g., fuzzy c-means clustering (Bezdek et al., 1984) or k-means clustering (Hartigan and Wong, 1979)), and Markov random field (MRF) based methods (Sun et al., 2021a). In this study, we directly use Otsu’s thresholding (Otsu, 1979) to extract the binary CM, and denote this as RIEM-O.

Second, we can replace the variable of change score vector \mathbf{p} in model (27) with the change label vector of \mathbf{L} , and solve the following energy model to obtain the final change label \mathbf{L}^* as

$$\mathbf{L}^* = \arg \min_{\mathbf{L} \in \{0,1\}^I} E_h^* (\mathbf{L}; \mathbf{X}, \mathbf{Y}, \alpha', \beta') \tag{32}$$

which is an integer quadratic optimization problem. Since the objective energy function is non-submodular (Gorelick et al., 2014; Jegelka and Bilmes, 2011; Kolmogorov and Rother, 2007), we cannot use traditional graph cut methods to solve the problem (32), such as the min-cut/max-flow method (Boykov and Kolmogorov, 2004). Here, we choose the local submodular approximation (LSA) based method (Gorelick et al., 2014) designed for minimizing non-submodular functions to obtain the change label \mathbf{L}^* , which iteratively approximates the energies locally. Similarly, we set the initial change label \mathbf{L}^0 to the Otsu segmented vector $\mathbf{B1} + \mathbf{B}^T \mathbf{1}$. Then, we compute the final CM as

$$\text{CM}(m, n) = L_i^*; \quad (m, n) \in A_i, \quad i = 1, \dots, N_S \tag{33}$$

We denote this CM calculation as RIEM-L. The framework of RIEM is summarized as in Algorithm 1.

Algorithm 1: Rules induced energy model for HCD.

Input: Heterogeneous images, parameters of N_S , α' , β' .

Structure representation:

Divide multitemporal images into superpixels.

Compute the feature distance matrices of \mathbf{D}^x and \mathbf{D}^y .

Construct the high-order KNN graphs of G_{hn}^{t1} and G_{hn}^{t2} .

Construct the high-order KFN graphs of G_{hf}^{t1} and G_{hf}^{t2} .

Energy model construction:

Compute the energy functions of $f_{i,j}^1$, $f_{i,j}^2$, $f_{i,j}^3$, $f_{i,j}^4$ and $\phi_{i,j}$.

Compute the matrices of \mathbf{B} and \mathbf{L}_w .

Construct the energy model E_h^* .

Change extraction:

Initialize the \mathbf{p}^0 and \mathbf{L}^0 .

Solve the energy minimization (29) to obtain \mathbf{p}^* .

Solve the energy minimization (32) to obtain \mathbf{L}^* .

Compute the DI by using (31) with \mathbf{p}^* .

Compute the CM by Otsu thresholding with \mathbf{p}^* or directly using \mathbf{L}^* .

4. Experiments and discussions

4.1. Experimental setting

4.1.1. Datasets

To evaluate the performance of RIEM, comparative experiments are conducted on seven real datasets,¹ as shown in Table 2 and Fig. 4. The heterogeneous pre-change and post-change images in these datasets are sourced from: (i) different optical sensors (i.e., SAR and optical sensors), such as Datasets #1–#4; (ii) sensors of different types, such as Datasets #5–#7. Additionally, these datasets cover various resolutions (30~0.52 m), image sizes (343 × 291~4135 × 2325), and different change events (e.g., floods, wildfires, land use changes), which can comprehensively assess the adaptability and robustness of different algorithms.

4.1.2. Evaluation metrics

We employ two categories of metrics to assess the obtained DI and CM. (i) Receiver Operating Characteristic (ROC) curves and Precision-Recall (PR) curves are utilized to assess the DI, with the area under ROC curve (AUR) and the area under PR curve (AUP) serving as quantitative metrics, respectively. (ii) Various colors are used to label true positive (TP), false positive (FP), true negative (TN), and false negative (FN) on the CM. Simultaneously, overall accuracy (OA), F1-score, Kappa coefficient (κ) and intersection over union (IoU) are employed as quantitative assessments for the CM.

4.1.3. Experimental settings

For all the experiment results of RIEM, we set the superpixel number $N_S = 2500$ for the employed GMMSP (Ban et al., 2018) (due to GMMSP cannot precisely control the number of segmented superpixels, the actual N_S is around 2500). Since the experimental datasets contains only SAR, RGB optical, and multispectral images, we chose the mean and median values as simple superpixel features to calculate feature distances of \mathbf{D}^x and \mathbf{D}^y (this is not exclusive, other features including those extracted by deep neural networks are also available). In addition, we fix the balancing parameters as $\alpha' = 15$ and $\beta' = 2^{-4}$ for the energy model $E_h^*(\mathbf{p}; \mathbf{X}, \mathbf{Y}, \alpha', \beta')$ of (29), and fix them as $\alpha' = 15$ and $\beta' = 1$ for the energy model $E_h^*(\mathbf{L}; \mathbf{X}, \mathbf{Y}, \alpha', \beta')$ of (32). These parameters will be discussed in Section 4.3.

¹ Datasets #1, #3, #4 and #7 (Mignotte, 2020) are available at <http://www-labs.iro.umontreal.ca/~mignotte>; Dataset #2 (Volpi et al., 2015) is available at <https://zenodo.org/records/8046719>; Dataset #5 (Luppino et al., 2019) is available at <https://sites.google.com/view/luppino/data>.

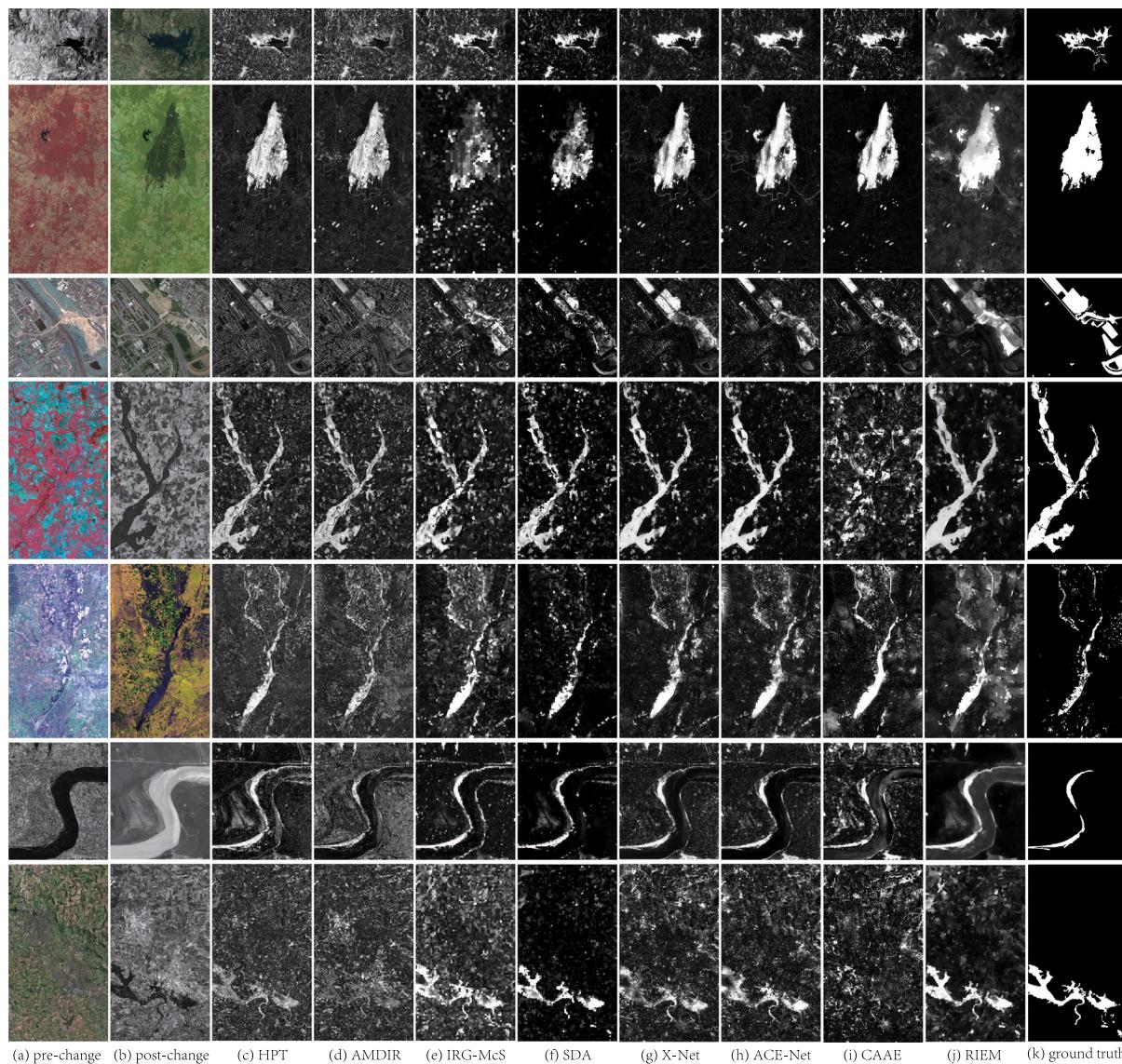


Fig. 4. Difference images on different datasets. From top to bottom, they are corresponding to Datasets #1 to #7, respectively. (a) pre-change image; (b) post-event image; from (c) to (j) are the difference images generated by (c) HPT, (d) AMDIR, (e) IRG-McS, (f) SDA, (g) X-Net, (h) ACE-Net, (i) CAAE, and (j) RIEM; (k) ground truth.

Table 2
Description of the heterogeneous datasets.

Dataset	Sensor (or modality)	Image size (pixels)	Date	Location	Spatial resolution	Change event
#1	Landsat-5 Google Earth	300 × 412 × 1 300 × 412 × 3	Sept. 1995 July 1996	Sardinia, Italy	30 m	Lake expansion
#2	Landsat-5 EO-1 ALI	1534 × 808 × 7 1534 × 808 × 10	Aug. 2011 Sept. 2011	Texas, USA	30 m	Forest fire
#3	Pleiades WorldView2	2000 × 2000 × 3 2000 × 2000 × 3	May 2012 July 2013	Toulouse, France	0.52 m	Construction
#4	NDVI SPOT	990 × 554 × 3 990 × 554 × 1	1999 2000	Gloucester, England	≈25 m	Flooding
#5	Landsat-8 Sentinel-1A	3500 × 2000 × 11 3500 × 2000 × 3	Jan. 2017 Feb. 2017	Sutter County, USA	≈15 m	Flooding
#6	Radarsat-2 Google Earth	343 × 291 × 1 343 × 291 × 3	June 2008 Sept. 2010	Yellow River, China	8 m	Embankment change
#7	QuickBird-2 TerraSAR-X	4135 × 2325 × 3 4135 × 2325 × 1	July 2006 July 2007	Gloucester, England	0.65 m	Flooding

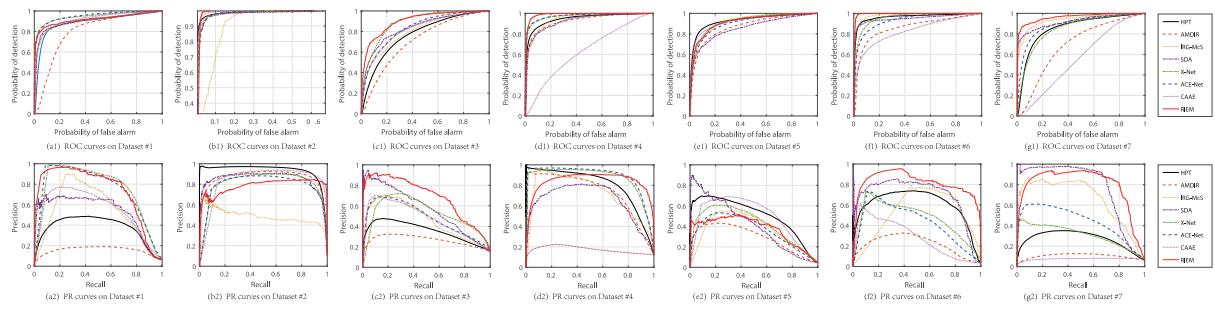


Fig. 5. ROC curves (top) and PR curves (bottom) of different methods. From (a) to (g) are the curves obtained from Datasets #1 to #7, respectively.

Table 3

AUR and AUP on different heterogeneous datasets. The highest scores are marked in bold.

Methods	Dataset #1		Dataset #2		Dataset #3		Dataset #4		Dataset #5		Dataset #6		Dataset #7		Average	
	AUR	AUP														
HPT (Liu et al., 2018b)	0.889	0.373	0.989	0.933	0.770	0.355	0.947	0.806	0.928	0.506	0.964	0.602	0.857	0.271	0.906	0.549
AMDIR (Luppino et al., 2019)	0.795	0.155	0.986	0.856	0.708	0.264	0.927	0.741	0.867	0.289	0.859	0.216	0.691	0.103	0.833	0.375
IRG-McS (Sun et al., 2021a)	0.899	0.520	0.924	0.481	0.805	0.460	0.946	0.711	0.918	0.358	0.981	0.513	0.937	0.656	0.916	0.528
SDA (Sun et al., 2022)	0.899	0.531	0.980	0.842	0.837	0.571	0.925	0.648	0.855	0.413	0.949	0.688	0.922	0.771	0.910	0.638
X-Net (Luppino et al., 2022)	0.946	0.737	0.984	0.798	0.885	0.520	0.973	0.865	0.918	0.412	0.957	0.465	0.851	0.309	0.931	0.586
ACE-Net (Luppino et al., 2022)	0.949	0.731	0.983	0.784	0.832	0.477	0.975	0.885	0.901	0.359	0.932	0.425	0.891	0.441	0.923	0.586
CAAE (Luppino et al., 2024)	0.917	0.540	0.992	0.831	0.818	0.468	0.632	0.174	0.921	0.482	0.828	0.278	0.559	0.075	0.809	0.407
Proposed RIEM	0.919	0.732	0.986	0.781	0.903	0.685	0.986	0.834	0.912	0.373	0.995	0.809	0.964	0.784	0.952	0.714

4.2. Experimental results

4.2.1. Difference images

To evaluate the DI of RIEM, we will compare RIEM with the following seven HCD methods.²

HPT (Liu et al., 2018b). A homogeneous pixel transformation based supervised method that first transforms one image to the other image domain by utilizing labeled samples (training set is established with 40% of unchanged pixels in the experiment), and then compares images in the same domain.

AMDIR (Luppino et al., 2019). A image regression based method that uses the affinity matrices distance (AMD) between heterogeneous images to pick possible unchanged pixels as a pseudo training set.

IRG-McS (Sun et al., 2021a). An unsupervised image structure comparison based method that first extract the structure feature of image by using a iterative robust graph, then compares the graphs and extract the changes by using an MRF based co-segmentation method.

SDA (Sun et al., 2022). An graph signal processing based method that transforms the pre-change image to the post-change image domain by utilizing a signal decomposition model based on the graph spectral domain analysis.

X-Net (Luppino et al., 2022). A deep translation network made up of two fully convolutional networks, which trains the network using the AMD-based change prior.

ACE-Net (Luppino et al., 2022). An adversarial cyclic encoder network that translates and compares the images, which also uses the AMD-based change prior.

CAAE (Luppino et al., 2024). A code-aligned autoencoders based method that aligns the latent code spaces of two autoencoders on affinity information extracted from heterogeneous images.

Fig. 4 shows the DIs obtained by RIEM and other comparison methods. Three facets can be noted: first, although most methods yield

DIs that can reflect certain change information, some methods are still not robust enough. For example, the DIs generated by AMDIR on Datasets #1 and #7, IRG-McS on Dataset #2, and CAAE on Datasets #4 and #7 are difficult to assess the change score of pixels. Second, the proposed RIEM consistently produces high-quality DIs across different datasets, where the discriminative ability between changed and unchanged regions is notable, validating the effectiveness and robustness of the rules induced energy model in HCD. Third, it is also observed that both RIEM and SDA yield sparse DIs as shown in Fig. 4(j) and (f), which is attributed to the incorporation of prior knowledge of change sparsity in the energy model and regression model, respectively. In this context, simple thresholding or clustering methods can effectively segment the changed and unchanged regions in the DIs generated by RIEM.

Fig. 5 plots the ROC and PR curves of the DI obtained by the different methods in Fig. 4, and the corresponding quantitative evaluation indicators of AUR and AUP are presented in Table 3. From Fig. 5 and Table 3, it can be seen that the DIs obtained by RIEM can well discriminate between changed and unchanged. RIEM obtains average AUR and AUP values of 0.952 and 0.714 respectively, which are 2.1% and 7.6% higher than those of the second-ranked (X-Net and SDA) respectively.

4.2.2. Change maps

To evaluate the CM of RIEM, we compare the proposed RIEM with the thirteen comparison methods, including not only the seven methods in Fig. 4 and Table 3, but also the following six methods.³

M3CD (Touati et al., 2020). An MRF based energy model that measures two pixels whether have identical labels or have different labels and employs the iterative conditional estimation in the energy minimization.

FPMS (Mignotte, 2020). An image translation based method that translate the pre-change image to the post-change image domain by utilizing the fractal encoding and fractal projection.

² HPT and AMDIR (under the “Image-Regression” subdirectory), X-Net and ACE-Net (under the “Deep-Image-Translation” subdirectory), and CAAE (under the “Code-Aligned-Autoencoders” subdirectory) are available at https://github.com/lulu025/Heterogeneous_CD; IRG-McS is available at <https://github.com/yulisun/IRG-McS>; SDA is available at <https://github.com/yulisun/HCD-GSP>.

³ M3CD, FPMS and CICM are available at <http://www-labs.iro.umontreal.ca/~mignotte/ResearchMaterial>; GSGM is available at <https://github.com/rshante0426/GSGM>; AOSG is available at <https://github.com/rshante0426/AOSG>; USSD is available at <https://github.com/zhuahu94854693/USSD>.

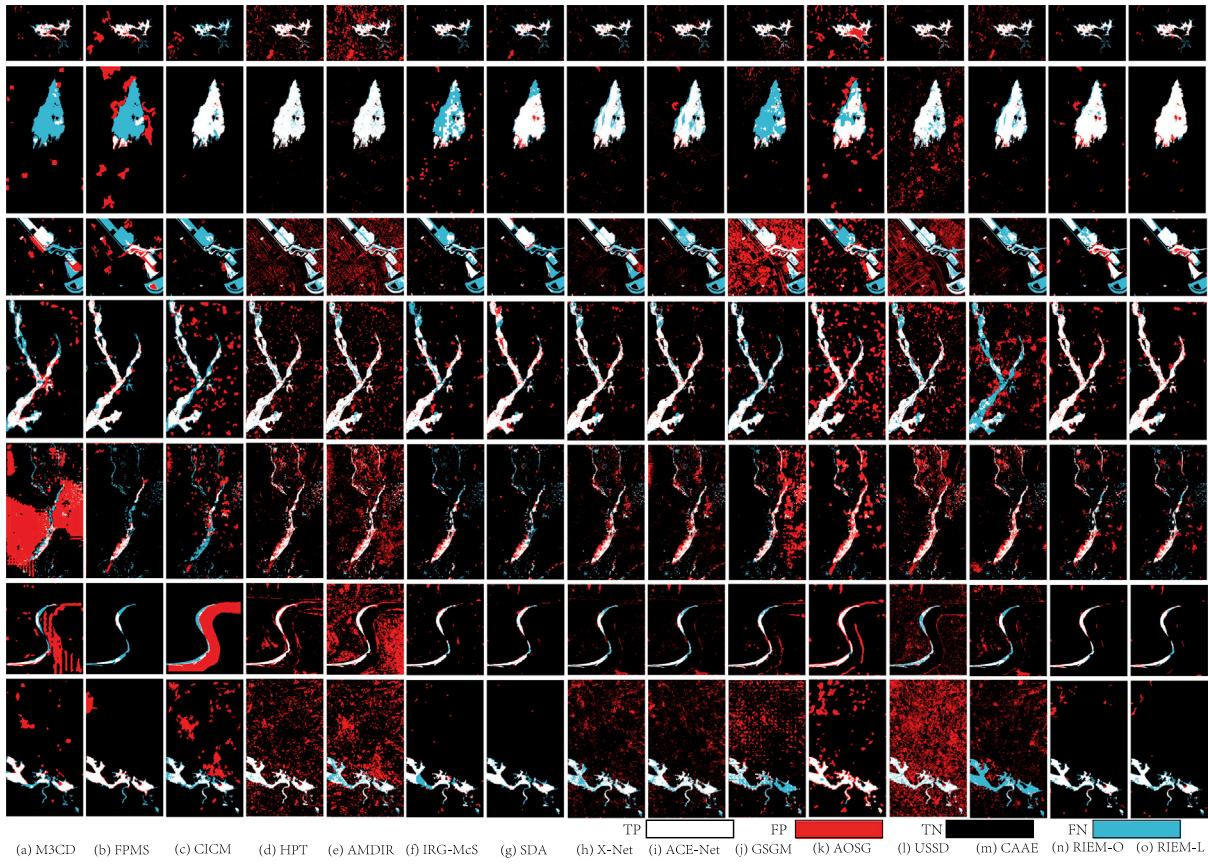


Fig. 6. Change maps generated by different methods. From top to bottom, they are corresponding to Datasets #1 to #7, respectively. From (a) to (o) are the change maps generated by (a) M3CD, (b) FPMS, (c) CICM, (d) HPT, (e) AMDIR, (f) IRG-McS, (g) SDA, (h) X-Net, (i) ACE-Net, (j) GSFM, (k) AOSG, (l) USSD, (m) CAAE, (n) RIEM-O, and (o) RIEM-L. In the change map, white, red, black, and cyan mark true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN), respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

CICM (Touati, 2019). A circular invariant convolution model based method that projects one image to the other image domain and then compares images by pixel differencing.

GSFM (Han et al., 2024b). A graph mapping method that constructs and compares the image structures represented by global structure graphs.

AOSG (Han et al., 2024c). An adaptive structured graph based method that cross-maps image structure features to the same domain to extract changes.

USSD (Zhu et al., 2024). An unsupervised spatial self-similarity difference based method that utilizes the image spatial relationships to detect changes.

The qualitative CMs obtained by different methods are presented in Fig. 6 and the corresponding quantitative evaluation indicators are presented in Table 4. From these results, it can be observed that HPT, AMDIR, AOSG and USSD have relatively more false positives, e.g., in Datasets #1, #3 and #7, while M3CD, FPMS, IRG-McS and GSFM have more false negatives in Dataset #2, resulting in smaller kappa coefficients. In Dataset #4, CAAE fails to detect the changes, causing a smaller F1-score of 0.239. The proposed RIEM outperforms other comparison methods on most datasets. For example, RIEM-L gains the highest F1-score on Datasets #1-#4, and RIEM-O gains the highest F1-score on Dataset #6. In terms of average scores (OA, κ , F1, IoU), RIEM-L and RIEM-O are ranked first and second respectively, and show significant improvement over the other methods, e.g., the average kappa coefficient of RIEM-L is 5.4% higher than the third-ranked SDA. This outstanding performance can be attributed to the following two aspects: first, RIEM is based on the universal six rules, which are derived from the nature of change detection problem without using

any assumptions, so RIEM is robust to different HCD scenarios and consistently achieves good performance with different datasets; second, RIEM takes full advantage of the flexibility of energy model, which not only takes into account the similarity and dissimilarity relationships of images, but also the spatial continuity and change sparsity, enabling the proposed model to more accurately portray the HCD problem. In addition, it can be observed that CMs of RIEM still contain some isolated false detections, especially in the RIEM-O. This is mainly due to the use of a simple thresholding strategy to generate binary CM, without incorporating any additional post-processing techniques such as morphological filtering (Chen et al., 2022). This also suggests that there is still room for further improvement in RIEM's performance through post-detection refinement.

4.3. Discussions

4.3.1. Parameter analysis

The main parameters used in RIEM are the superpixel number of N_S , the weighting parameters of α' and β' in the energy model.

In general, the superpixel number, N_S , should be determined according to the spatial resolution of the dataset and the timeliness requirement of the HCD task. A larger N_S can result in smaller segmented superpixel areas, thereby improving the granularity of change detection. Fig. 7 shows the DIs and CMs generated by RIEM on Datasets #3 and #5 with different N_S , where we mark some details with the elliptical regions. It can be found that more details can be detected as N_S increases. However, when we choose a larger N_S , the computational complexity of the algorithm also rises, as discussed later in this subsection. Conversely, reducing the N_S can enhance computational

Table 4

OA, κ , F1 and IoU of CMs. The best and second best scores are marked in red and blue, respectively.

Methods	Dataset #1				Dataset #2				Dataset #3				Dataset #4				Dataset #5				Dataset #6				Dataset #7				Average			
	OA	κ	F1	IoU																												
M3CD (Touati et al., 2020)	0.956	0.633	0.656	0.488	0.884	0.044	0.075	0.039	0.863	0.406	0.482	0.318	0.915	0.588	0.636	0.466	0.575	0.021	0.077	0.040	0.856	0.158	0.204	0.113	0.952	0.618	0.643	0.474	0.857	0.352	0.396	0.277
FPMS (Mignotte, 2020)	0.938	0.593	0.625	0.454	0.830	0.019	0.112	0.059	0.827	0.269	0.368	0.226	0.962	0.816	0.837	0.720	0.947	0.329	0.356	0.217	0.979	0.544	0.553	0.382	0.970	0.770	0.786	0.648	0.922	0.477	0.520	0.387
CICM (Touati, 2019)	0.943	0.451	0.481	0.317	0.983	0.909	0.918	0.849	0.867	0.270	0.321	0.191	0.884	0.507	0.573	0.401	0.899	0.081	0.131	0.070	0.789	0.024	0.080	0.042	0.896	0.371	0.423	0.268	0.894	0.373	0.418	0.305
HPT (Liu et al., 2018b)	0.912	0.504	0.546	0.376	0.987	0.930	0.937	0.881	0.815	0.415	0.523	0.354	0.918	0.671	0.717	0.558	0.932	0.488	0.518	0.350	0.927	0.447	0.476	0.312	0.860	0.397	0.455	0.294	0.907	0.550	0.596	0.447
AMDIR (Luppino et al., 2019)	0.799	0.255	0.328	0.196	0.982	0.904	0.914	0.842	0.724	0.259	0.411	0.258	0.898	0.607	0.664	0.497	0.822	0.236	0.291	0.170	0.686	0.116	0.170	0.093	0.782	0.171	0.254	0.145	0.813	0.364	0.433	0.315
IRG-McS (Sun et al., 2021a)	0.971	0.739	0.754	0.605	0.912	0.448	0.493	0.327	0.882	0.420	0.478	0.314	0.939	0.714	0.749	0.599	0.959	0.490	0.512	0.344	0.976	0.690	0.702	0.541	0.971	0.740	0.755	0.606	0.944	0.606	0.635	0.477
SDA (Sun et al., 2022)	0.958	0.654	0.676	0.511	0.973	0.856	0.871	0.771	0.902	0.523	0.572	0.400	0.947	0.768	0.798	0.664	0.954	0.433	0.457	0.296	0.981	0.728	0.738	0.585	0.979	0.806	0.817	0.690	0.956	0.681	0.704	0.560
X-Net (Luppino et al., 2022)	0.967	0.723	0.740	0.587	0.966	0.815	0.834	0.715	0.874	0.471	0.544	0.374	0.947	0.768	0.799	0.664	0.931	0.428	0.461	0.300	0.962	0.526	0.546	0.375	0.878	0.341	0.401	0.251	0.932	0.582	0.618	0.467
ACE-Net (Luppino et al., 2022)	0.958	0.674	0.696	0.534	0.961	0.790	0.811	0.683	0.870	0.397	0.466	0.304	0.952	0.789	0.816	0.690	0.911	0.367	0.406	0.255	0.959	0.490	0.510	0.343	0.901	0.447	0.497	0.330	0.930	0.565	0.600	0.448
GSGM (Han et al., 2024b)	0.958	0.677	0.699	0.538	0.901	0.255	0.294	0.172	0.519	0.074	0.292	0.171	0.935	0.684	0.720	0.563	0.858	0.233	0.286	0.167	0.963	0.575	0.593	0.422	0.851	0.165	0.237	0.134	0.855	0.381	0.446	0.310
AOSG (Han et al., 2024c)	0.874	0.397	0.451	0.291	0.916	0.589	0.636	0.466	0.782	0.216	0.346	0.209	0.842	0.486	0.569	0.397	0.898	0.329	0.372	0.228	0.942	0.502	0.526	0.357	0.910	0.532	0.574	0.402	0.881	0.436	0.496	0.336
USSD (Zhu et al., 2024)	0.965	0.726	0.745	0.593	0.916	0.632	0.678	0.513	0.694	0.214	0.379	0.234	0.878	0.528	0.597	0.425	0.896	0.384	0.425	0.270	0.880	0.152	0.196	0.109	0.578	0.084	0.187	0.103	0.829	0.389	0.458	0.321
CAAE (Luppino et al., 2024)	0.930	0.534	0.569	0.398	0.974	0.861	0.875	0.777	0.867	0.380	0.451	0.291	0.803	0.126	0.239	0.136	0.930	0.456	0.488	0.323	0.933	0.324	0.355	0.215	0.844	0.014	0.069	0.036	0.897	0.385	0.435	0.311
RIEM-O (proposed)	0.971	0.730	0.745	0.594	0.973	0.863	0.878	0.783	0.902	0.577	0.633	0.463	0.947	0.780	0.810	0.681	0.940	0.463	0.492	0.327	0.985	0.796	0.803	0.671	0.974	0.785	0.799	0.665	0.956	0.713	0.737	0.598
RIEM-L (proposed)	0.970	0.744	0.760	0.613	0.987	0.935	0.942	0.890	0.923	0.669	0.713	0.553	0.958	0.817	0.841	0.725	0.954	0.473	0.499	0.332	0.981	0.732	0.742	0.589	0.976	0.779	0.792	0.656	0.963	0.735	0.755	0.623

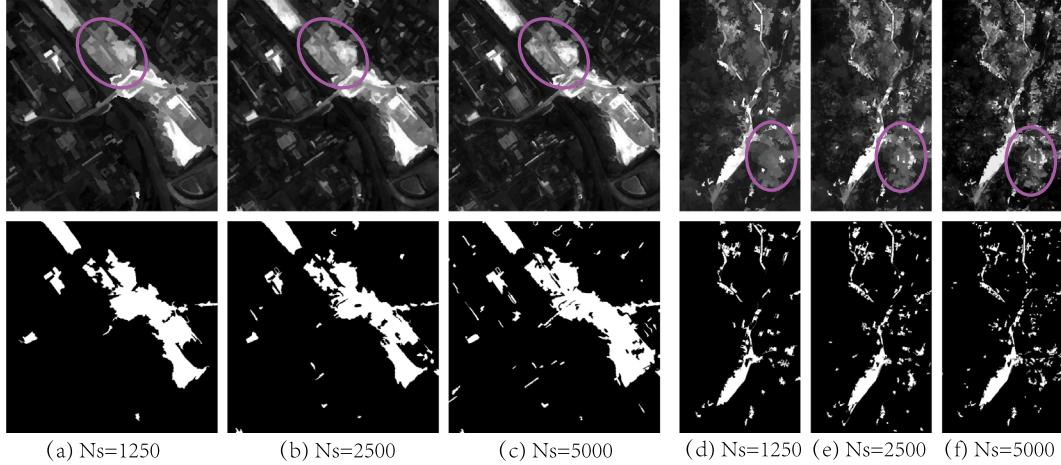


Fig. 7. Difference images (top row) and change maps (bottom row) of RIEM on Datasets #3 and #5 with different N_s . (a)–(c) are the results on Dataset #3 under $N_s = 1250, 2500, 5000$, the corresponding κ values of CMs are 0.558, 0.577, and 0.580, respectively. (d)–(f) are the results on Dataset #5 under $N_s = 1250, 5000$, the corresponding κ values of CMs are 0.443, 0.463, and 0.469, respectively.

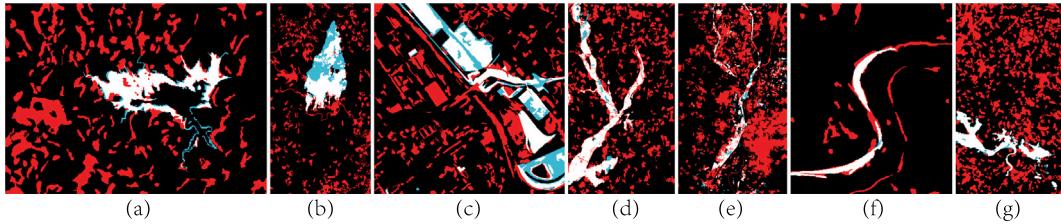


Fig. 8. Change maps obtained by RIEM-O with $\alpha' = 0$. From (a) to (g) are the results on Datasets #1 to #7.

efficiency, but if the N_s is too small, the segmented superpixel areas may become excessively large, potentially causing some superpixels to encompass different land cover categories, i.e., disrupting the internal homogeneity of superpixels, which may degrade the detection accuracy as shown in Fig. 7. As a compromise, this paper simply sets $N_s \approx 2500$.

For the weighting parameters of α' and β' , they are applied to adjust the weights of the smooth and sparse constraint terms in the energy function (27), respectively. First, we conduct the ablation analysis. (1) If the RIEM model focuses only on the similarity/dissimilarity relationships consistency term of $(\mathbf{1} - \mathbf{p})^T \mathbf{B} (\mathbf{1} - \mathbf{p})$, i.e., set $\alpha' \rightarrow 0$ and $\beta' \rightarrow 0$, the solution of (27) will be $\mathbf{p}^* = \mathbf{1}$ with $E_h^* = 0$; and if the RIEM model remove this relationships consistency term, the solution of (27) will be $\mathbf{p}^* = \mathbf{0}$. (2) If the RIEM model focuses only on the smooth constraint term of $\mathbf{p}^T \mathbf{L}_{wp} \mathbf{p}$, i.e., set $\alpha' \rightarrow \infty$, the solution of (27) will be $\mathbf{p}^* = \mathbf{c1}$ with c being a constant. Fig. 8 shows the CMs obtained by RIEM-O with $\alpha' = 0$, from which it can be seen that there are considerable salt-and-pepper noise in CM when the energy model discards the smooth constraint. (3) If the RIEM model focuses only on the sparse constraint term, i.e., set $\beta' \rightarrow \infty$, the solution of (27) will be $\mathbf{p}^* = \mathbf{0}$; and if the RIEM model remove the sparse constraint term, the solution of (27) will be $\mathbf{p}^* = \mathbf{1}$.

Second, we show parameters sensitivity of RIEM in Fig. 9. To avoid redundancy, we only show the κ of RIEM-L with different α' (from 10 to 20) and RIEM-O with different β' (from 2^{-6} to 2^{-1}). From Fig. 9 it can be seen that RIEM gains favorable performance to a certain range of values of α' and β' , which shows the robustness to these weighting parameters. In addition, the reason for setting different values of β' in RIEM-L and RIEM-O is that β' is used to balance the first-power and second-power functions regarding to \mathbf{L} and \mathbf{p} , respectively. In the latter, \mathbf{p} is taken as a decimal between 0 and 1, causing more attenuation of the energy functions with respect to the second-power of \mathbf{p} (i.e., the relationships consistency term and smooth constraint term). Therefore, a smaller β' is needed in the RIEM-O to reduce the weights of the first-power loss functions (i.e., the sparse constraint term).

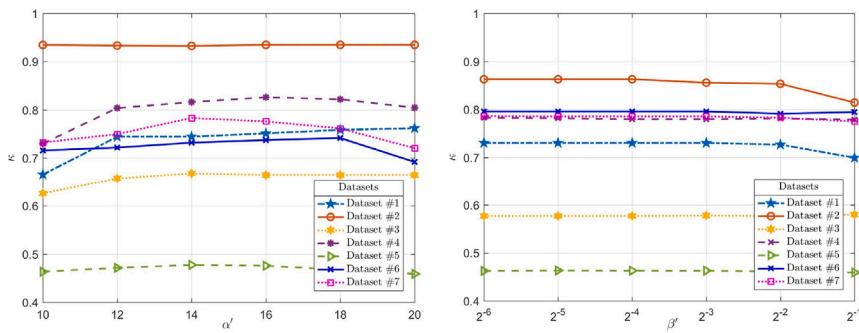
Table 5

Ablation study of the four energy functions in RIEM measured by the average scores.

Methods	Energy functions in E_h				Performance			
	E_{sr}	E_{dr}	E_{ss}	E_{ps}	OA	κ	F1	IoU
Settings	–	✓	–	✓	0.899	0.427	0.475	0.319
	–	✓	✓	✓	0.928	0.537	0.574	0.408
	✓	–	–	✓	0.934	0.618	0.653	0.495
	✓	✓	–	✓	0.958	0.696	0.719	0.575
RIEM	✓	–	✓	✓	0.958	0.700	0.724	0.581
	✓	✓	✓	✓	0.963	0.735	0.755	0.623

4.3.2. Ablation study

The hybrid energy function E_h (24) in RIEM consists of four components: E_{sr} , E_{dr} , E_{ss} , and E_{ps} . From the expression of E_h , it can be observed that each energy term exhibits a specific tendency: E_{sr} and E_{dr} both encourage $\mathbf{p} = \mathbf{1}$, E_{ss} favors $\mathbf{p} = \mathbf{c1}$ with c being a constant, and E_{ps} tends towards $\mathbf{p} = \mathbf{0}$. Therefore, using only a subset of the energy terms may lead to trivial solutions. For example, combining $E_{sr} + E_{dr}$, $E_{sr} + E_{ss}$, $E_{dr} + E_{ss}$, $E_{sr} + E_{dr} + E_{ss}$ would result in the trivial solution of $\mathbf{p}^* = \mathbf{1}$, while $E_{ss} + E_{ps}$ would yield $\mathbf{p}^* = \mathbf{0}$. In addition to these combinations that lead to trivial results, we evaluate the performance of the remaining six valid energy combinations, using the average OA, κ , F1 and IoU as evaluation metrics. As shown in Table 5, it is clear that the performance of RIEM degrades when any one of E_{sr} , E_{dr} or E_{ss} is removed. Specifically, removing the similarity relationships based energy E_{sr} (derived from Rules 1–4) leads to the largest decrease in F1-score and IoU, highlighting its crucial role in exploiting intra-image similarity for change detection. The dissimilarity relationships based energy E_{dr} (derived from Rules 5–6) and the spatial smoothness based energy E_{ss} also show significant impact: E_{dr} can

Fig. 9. Sensitivity analysis of parameters α' (left) and β' (right).**Table 6**The computational time (s) of different methods on Dataset #3 with image size $2000 \times 2000 \times 3$.

Methods	M3CD (Touati et al., 2020)	FPMS (Mignotte, 2020)	CICM (Touati, 2019)	HPT (Liu et al., 2018b)	AMDIR (Luppino et al., 2019)	IRG-McS (Sun et al., 2021a)	SDA (Sun et al., 2022)	GSGM (Han et al., 2024b)	AOSG (Han et al., 2024c)	USSD (Zhu et al., 2024)	RIEM-O	RIEM-L
Time (s)	2611.6	92.4	113.0	78.1	109.6	20.5	10.2	788.8	207.3	795.2	5.5	5.2

effectively utilize intra-image dissimilarity relationships, while E_{ss} can mitigate the salt-and-pepper noise in CM by incorporating both spatial context and similarity/dissimilarity relationships in multitemporal images. Overall, these results confirm that all four energy components are complementary and jointly contribute to the superior performance of the proposed RIEM.

4.3.3. Complexity analysis

The proposed RIEM consists of three processes: preprocessing, constructing the energy model and solving the energy model. In the preprocessing, the complexity of superpixel segmentation is $\mathcal{O}(MN)$ as illustrated in Ban et al. (2018). In the energy model construction, computing the feature distance matrices of \mathbf{D}^x and \mathbf{D}^y needs $\mathcal{O}((B_x + B_y) N_S^2)$; sorting the feature distance matrix and construct the high-order KNN and KFN graphs needs $\mathcal{O}(N_S^2 \log N_S)$; computing the matrix \mathbf{B} needs $\mathcal{O}(N_S^2)$; computing the matrix \mathbf{L}_w needs $\mathcal{O}(N_S^2 + N_R)$ with $N_R = \sum_{i \in I} |\mathcal{N}_i^R|$. In the energy model solution, for the RIEM-O with (29), the gradient descent method requires a complexity of $\mathcal{O}(N_S^2)$ to compute the gradient vector per iteration; for the RIEM-L with (32), the LSA method is used for this non-submodular energy minimization, whose complexity is presented in Gorelick et al. (2014). Generally speaking, the complexity of RIEM is relatively low, and it is mainly related to the number of superpixel N_S rather than the number of image pixels.

Table 6 lists the running times of the compared traditional algorithms on Dataset #3 with their original open source codes and default parameters, where the C++ codes of M3CD, FPMS, and CICM are executed in a Linux environment with Intel Xeon Silver 4110 CPU, HPT, AMDIR, IRG-McS, SDA, GSGM, AOSG, USSD and RIEM are implemented in MATLAB and ran on a Windows environment with Intel Core i7-8700K CPU. As observed from Table 6, the proposed superpixel-based RIEM has the least runtime, demonstrating the high efficiency and significant practical applicability.

5. Conclusion

This paper proposes a rules-induced unsupervised HCD method. Starting from the nature of change detection problem, it first establishes a general connection between pairwise relationships and change labels, and accordingly derives six universal rules for determining the change labels. These labels provide a nice solution for HCD problem: detecting changes by considering only the intra-image relationships within each image rather than considering the inter-image comparisons, which significantly reduces the difficulty of HCD problem. Based on this, it

constructs an energy model to portray these rules, and eventually solves the rules-induced energy model to directly extract the change information. Experimental results on seven datasets, compared with thirteen representative methods, demonstrate the effectiveness of the proposed approach. RIEM achieves an average improvement of 5.1% in F1-score while maintaining the lowest runtime. These results confirm the feasibility, superiority, and practical potential of RIEM for real-world HCD applications. Given the favorable properties of the derived rules, we hope they can inspire a rethinking of the change detection problem, especially as deep learning based methods are being systematically applied.

CRediT authorship contribution statement

Yuli Sun: Writing – original draft, Software, Methodology, Investigation. **Lin Lei:** Writing – review & editing, Investigation, Formal analysis. **Zhang Li:** Writing – review & editing, Visualization, Investigation. **Gangyao Kuang:** Supervision, Project administration. **Qifeng Yu:** Writing – review & editing, Visualization, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors would like to thank Prof. Max Mignotte, Prof. Michele Volpi, Dr. Redha Touati and Dr. Luigi Tommaso Luppino for sharing the datasets and codes. This work was supported by the National Natural Science Foundation of China [grant number 62401577]; the Natural Science Foundation of Hunan Province of China [grant number 2024JJ6466]; the China Postdoctoral Science Foundation, China [grant number 2024M754301]; and the Postdoctoral Fellowship Program of China Postdoctoral Science Foundation [grant number GZC20233545].

References

- Agoritsas, E., Catania, G., Decelle, A., Seoane, B., 2023. Explaining the effects of non-convergent MCMC in the training of energy-based models. In: Krause, A., Brunskill, E., Cho, K., Engelhardt, B., Sabato, S., Scarlett, J. (Eds.), Proceedings of the 40th International Conference on Machine Learning. In: Proceedings of Machine Learning Research, Vol. 202, PMLR, pp. 322–336, [Online]. Available: <https://proceedings.mlr.press/v202/agoritsas23a.html>.
- Ban, Z., Liu, J., Cao, L., 2018. Superpixel segmentation using Gaussian mixture model. IEEE Trans. Image Process. 27 (8), 4105–4117.

- Bezdek, J.C., Ehrlich, R., Full, W., 1984. FCM: The fuzzy c-means clustering algorithm. *Comput. Geosci.* 10 (2–3), 191–203.
- Boykov, Y., Kolmogorov, V., 2004. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE Trans. Pattern Anal. Mach. Intell.* 26 (9), 1124–1137.
- Camps-Valls, G., Gomez-Chova, L., Munoz-Mari, J., Rojo-Alvarez, J.L., Martinez-Ramon, M., 2008. Kernel-Based framework for multitemporal and multisource remote sensing data classification and change detection. *IEEE Trans. Geosci. Remote Sens.* 46 (6), 1822–1835.
- Cartwright, D., Harary, F., 1956. Structural balance: a generalization of heider's theory. *Psychol Rev* 63 (5), 277.
- Chen, Y., Bruzzone, L., 2022. Self-Supervised change detection in multiview remote sensing images. *IEEE Trans. Geosci. Remote Sens.* 60, 1–12.
- Chen, J., Hou, D., He, C., Liu, Y., Guo, Y., Yang, B., 2024a. Change detection with cross-domain remote sensing images: A systematic review. *IEEE J. Sel. Top. Appl. Earth Obs. Remote. Sens.* 17, 11563–11582.
- Chen, H., Lan, C., Song, J., Broni-Bediako, C., Xia, J., Yokoya, N., 2024b. ObjFormer: Learning land-cover changes from paired OSM data and optical high-resolution imagery via Object-Guided transformer. *IEEE Trans. Geosci. Remote Sens.* 62, 1–22.
- Chen, H., Song, J., Wu, C., Du, B., Yokoya, N., 2023a. Exchange means change: An unsupervised single-temporal change detection framework based on intra- and inter-image patch exchange. *ISPRS J. Photogramm. Remote Sens.* 206, 87–105, [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S092427162300309X>.
- Chen, H., Yokoya, N., Chini, M., 2023b. Fourier domain structural relationship analysis for unsupervised multimodal change detection. *ISPRS J. Photogramm. Remote Sens.* 198, 99–114.
- Chen, H., Yokoya, N., Wu, C., Du, B., 2022. Unsupervised multimodal change detection based on structural relationship graph representation learning. *IEEE Trans. Geosci. Remote Sens.* 60, 1–18.
- Dai, A., Yang, J., Zhang, T., Gao, B., Tang, K., Chen, X., 2024. DADR-HCD: A deep domain adaptation and disentangled representation network for unsupervised heterogeneous change detection. *IEEE Trans. Geosci. Remote Sens.* 62, 1–15.
- Deledalle, C.-A., Denis, L., Tupin, F., 2012. How to compare noisy patches? Patch similarity beyond Gaussian noise. *Int. J. Comput. Vis.* 99, 86–102.
- Delong, A., Osokin, A., Isack, H.N., Boykov, Y., 2012. Fast approximate energy minimization with label costs. *Int. J. Comput. Vis.* 96, 1–27.
- Florez-Ospina, J.F., Jimenez-Sierra, D.A., Benitez-Restrepo, H.D., Arce, G.R., 2023. Exploiting variational inequalities for generalized change detection on graphs. *IEEE Trans. Geosci. Remote Sens.* 61, 1–16.
- Fukunaga, K., Hostetler, L., 1973. Optimization of k nearest neighbor density estimates. *IEEE Trans. Inform. Theory* 19 (3), 320–326.
- Gorelick, L., Boykov, Y., Veksler, O., Ben Ayed, I., Delong, A., 2014. Submodularization for binary pairwise energies. In: CVPR. pp. 1154–1161.
- Han, T., Tang, Y., Chen, Y., Yang, X., Guo, Y., Jiang, S., 2024a. SDC-GAE: Structural difference compensation graph autoencoder for unsupervised multimodal change detection. *IEEE Trans. Geosci. Remote Sens.* 62, 1–16.
- Han, T., Tang, Y., Chen, Y., Zou, B., Feng, H., 2024b. Global structure graph mapping for multimodal change detection. *Int. J. Digit. Earth* 17 (1), 2347457.
- Han, T., Tang, Y., Zou, B., Feng, H., 2024c. Unsupervised multimodal change detection based on adaptive optimization of structured graph. *Int. J. Appl. Earth Obs. Geoinf.* 126, 103630, [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1569843223004545>.
- Hartigan, J.A., Wong, M.A., 1979. Algorithm AS 136: A k-means clustering algorithm. *J. R. Stat. Soc. Ser. C (Appl. Stat.)* 28 (1), 100–108.
- Heider, F., 1946. Attitudes and cognitive organization. *J. Psychol.* 21 (1), 107–112.
- Hong, D., Zhang, B., Li, X., Li, Y., Li, C., Yao, J., Yokoya, N., Li, H., Ghamisi, P., Jia, X., Plaza, A., Gamba, P., Benediktsson, J.A., Chanussot, J., 2024. SpectralGPT: Spectral remote sensing foundation model. *IEEE Trans. Pattern Anal. Mach. Intell.* 46 (8), 5227–5244.
- Isack, H., Boykov, Y., 2012. Energy-based geometric multi-model fitting. *Int. J. Comput. Vis.* 97 (2), 123–147.
- Jegelka, S., Bilmes, J., 2011. Submodularity beyond submodular energies: Coupling edges in graph cuts. In: CVPR. pp. 1897–1904.
- Jia, M., Zhang, C., Lv, Z., Zhao, Z., Wang, L., 2022a. Bipartite adversarial autoencoders with structural self-similarity for unsupervised heterogeneous remote sensing image change detection. *IEEE Geosci. Remote. Sens. Lett.* 19, 1–5.
- Jia, M., Zhang, C., Zhao, Z., Wang, L., 2022b. Bipartite graph attention autoencoders for unsupervised change detection using VHR remote sensing images. *IEEE Trans. Geosci. Remote Sens.* 60, 1–15.
- Kolmogorov, V., Rother, C., 2007. Minimizing nonsubmodular functions with graph Cuts-A review. *IEEE Trans. Pattern Anal. Mach. Intell.* 29 (7), 1274–1279.
- LeCun, Y., Chopra, S., Hadsell, R., Ranzato, M., Huang, F., 2006. A tutorial on energy-based learning. *Predict. Struct. Data* 1.
- Li, X., Du, Z., Huang, Y., Tan, Z., 2021. A deep translation (GAN) based change detection network for optical and SAR remote sensing images. *ISPRS J. Photogramm. Remote Sens.* 179, 14–34.
- Li, C., Li, G., Wang, Z., Wang, X., Varshney, P.K., 2024. COMIC: An unsupervised change detection method for heterogeneous remote sensing images based on copula mixtures and Cycle-Consistent adversarial networks. *Inf. Fusion* 102240, [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1566253524000186>.
- Liu, J., Gong, M., Qin, K., Zhang, P., 2018a. A deep convolutional coupling network for change detection based on heterogeneous optical and radar images. *IEEE Trans. Neural Netw. Learn. Syst.* 29 (3), 545–559.
- Liu, Z., Li, G., Mercier, G., He, Y., Pan, Q., 2018b. Change detection in heterogeneous remote sensing images via homogeneous pixel transformation. *IEEE Trans. Image Process.* 27 (4), 1822–1834.
- Liu, Z.-g., Mercier, G., Dezert, J., Pan, Q., 2014. Change detection in heterogeneous remote sensing images based on multidimensional evidential reasoning. *IEEE Geosci. Remote. Sens. Lett.* 11 (1), 168–172.
- Loftsgaarden, D.O., Quesenberry, C.P., 1965. A nonparametric estimate of a multivariate density function. *Ann. Math. Stat.* 36 (3), 1049–1051.
- Luppino, L.T., 2020. Unsupervised Change Detection in Heterogeneous Remote Sensing Imagery (Ph.D. dissertation). UiT The Arctic University of Norway.
- Luppino, L.T., Bianchi, F.M., Moser, G., Anfinsen, S.N., 2019. Unsupervised image regression for heterogeneous change detection. *IEEE Trans. Geosci. Remote Sens.* 57 (12), 9960–9975.
- Luppino, L.T., Hansen, M.A., Kampffmeyer, M., Bianchi, F.M., Moser, G., Jenssen, R., Anfinsen, S.N., 2024. Code-Aligned autoencoders for unsupervised change detection in multimodal remote sensing images. *IEEE Trans. Neural Netw. Learn. Syst.* 35 (1), 60–72.
- Luppino, L.T., Kampffmeyer, M., Bianchi, F.M., Moser, G., Serpico, S.B., Jenssen, R., Anfinsen, S.N., 2022. Deep image translation with an Affinity-Based change prior for unsupervised multimodal change detection. *IEEE Trans. Geosci. Remote Sens.* 60, 1–22.
- Lv, Z., Huang, H., Gao, L., Benediktsson, J.A., Zhao, M., Shi, C., 2022a. Simple multiscale UNet for change detection with heterogeneous remote sensing images. *IEEE Geosci. Remote. Sens. Lett.* 19, 1–5.
- Lv, Z., Huang, H., Li, X., Zhao, M., Benediktsson, J.A., Sun, W., Falco, N., 2022b. Land cover change detection with heterogeneous remote sensing images: Review, progress, and perspective. *Proc. IEEE* 110 (12), 1976–1991.
- Lv, Z., Huang, H., Sun, W., Jia, M., Benediktsson, J.A., Chen, F., 2024. Iterative training sample augmentation for enhancing land cover change detection performance with deep learning neural network. *IEEE Trans. Neural Netw. Learn. Syst.* 1–14.
- Lv, Z., Huang, H., Sun, W., Lei, T., Benediktsson, J.A., Li, J., 2023. Novel enhanced UNet for change detection using multimodal remote sensing image. *IEEE Geosci. Remote. Sens. Lett.* 20, 1–5.
- McGregor, I.R., Connette, G., Gray, J.M., 2024. A multi-source change detection algorithm supporting user customization and near real-time deforestation detections. *Remote Sens. Environ.* 308, 114195.
- Mercier, G., Moser, G., Serpico, S.B., 2008. Conditional copulas for change detection in heterogeneous remote sensing images. *IEEE Trans. Geosci. Remote Sens.* 46 (5), 1428–1441.
- Mignotte, M., 2020. A fractal projection and Markovian segmentation-based approach for multimodal change detection. *IEEE Trans. Geosci. Remote Sens.* 58 (11), 8046–8058.
- Mitra, P., Murthy, C., Pal, S.K., 2002. Unsupervised feature selection using feature similarity. *IEEE Trans. Pattern Anal. Mach. Intell.* 24 (3), 301–312.
- Otsu, N., 1979. A threshold selection method from gray-level histograms. *IEEE Trans. Syst. Man Cybern.* 9 (1), 62–66.
- Shi, J., Wu, T., Kai Qin, A., Lei, Y., Jeon, G., 2024. Self-Guided autoencoders for unsupervised change detection in heterogeneous remote sensing images. *IEEE Trans. Artif. Intell.* 5 (6), 2458–2471.
- Sun, Y., Lei, L., Guan, D., Kuang, G., 2021a. Iterative robust graph for unsupervised change detection of heterogeneous remote sensing images. *IEEE Trans. Image Process.* 30, 6277–6291.
- Sun, Y., Lei, L., Guan, D., Kuang, G., Li, Z., Liu, L., 2024a. Locality preservation for unsupervised multimodal change detection in remote sensing imagery. *IEEE Trans. Neural Netw. Learn. Syst.* 1–15.
- Sun, Y., Lei, L., Guan, D., Kuang, G., Liu, L., 2022. Graph signal processing for heterogeneous change detection. *IEEE Trans. Geosci. Remote Sens.* 60, 1–23.
- Sun, Y., Lei, L., Li, Z., Kuang, G., 2024b. Similarity and dissimilarity relationships based graphs for multimodal change detection. *ISPRS J. Photogramm. Remote Sens.* 208, 70–88, [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0924271624000029>.
- Sun, Y., Lei, L., Li, X., Tan, X., Kuang, G., 2021b. Structure consistency-based graph for unsupervised change detection with homogeneous and heterogeneous remote sensing images. *IEEE Trans. Geosci. Remote Sens.* 1–21.
- Sun, Y., Lei, L., Liu, L., Kuang, G., 2023. Structural regression fusion for unsupervised multimodal change detection. *IEEE Trans. Geosci. Remote Sens.* 61, 1–18.
- Touati, R., 2019. Détection de Changement en Imagerie Satellitaire Multimodale (Ph.D. dissertation). Université de Montréal.
- Touati, R., Mignotte, M., 2018. An Energy-Based model encoding nonlocal pairwise pixel interactions for multisensor change detection. *IEEE Trans. Geosci. Remote Sens.* 56 (2), 1046–1058.

- Touati, R., Mignotte, M., Dahmane, M., 2020. Multimodal change detection in remote sensing images using an unsupervised pixel pairwise-based Markov random field model. *IEEE Trans. Image Process.* 29, 757–767.
- Volpi, M., Camps-Valls, G., Tuia, D., 2015. Spectral alignment of multi-temporal cross-sensor images with automated kernel canonical correlation analysis. *ISPRS J. Photogramm. Remote Sens.* 107, 50–63.
- Wan, L., Xiang, Y., You, H., 2019. An object-based hierarchical compound classification method for change detection in heterogeneous optical and SAR images. *IEEE Trans. Geosci. Remote Sens.* 57 (12), 9941–9959.
- Wang, J.-J., Dobigeon, N., Chabert, M., Wang, D.-C., Huang, T.-Z., Huang, J., 2024. CD-GAN: A robust fusion-based generative adversarial network for unsupervised remote sensing change detection with heterogeneous sensors. *Inf. Fusion* 102313.
- Wu, C., Du, B., Zhang, L., 2023. Fully convolutional change detection framework with generative adversarial network for unsupervised, weakly supervised and regional supervised change detection. *IEEE Trans. Pattern Anal. Mach. Intell.* 45 (8), 9774–9788.
- Wu, H., Geng, J., Jiang, W., 2024a. Multidomain constrained translation network for change detection in heterogeneous remote sensing images. *IEEE Trans. Geosci. Remote Sens.* 62, 1–16.
- Wu, Y., Li, J., Yuan, Y., Qin, A.K., Miao, Q.-G., Gong, M.-G., 2021. Commonality autoencoder: Learning common features for change detection from heterogeneous images. *IEEE Trans. Neural Netw. Learn. Syst.* 1–14.
- Wu, C., Zhang, L., Du, B., Chen, H., Wang, J., Zhong, H., 2024b. UNet-Like remote sensing change detection: A review of current models and research directions. *IEEE Geosci. Remote. Sens. Mag.* 2–31.
- Xing, Y., Zhang, Q., Ran, L., Zhang, X., Yin, H., Zhang, Y., 2024. Improving reliability of heterogeneous change detection by sample synthesis and knowledge transfer. *IEEE Trans. Geosci. Remote Sens.* 62, 1–11.
- Yu, Y., Shen, Y., Liu, Y., Wei, Y., Rui, X., Li, B., 2023. Knowledge mapping and trends in research on remote sensing change detection using CiteSpace analysis. *Earth Sci. Inform.* 16 (1), 787–801.
- Zhao, B., Sui, H., Liu, J., Shi, W., Wang, W., Xu, C., Wang, J., 2024. Flood inundation monitoring using multi-source satellite imagery: a knowledge transfer strategy for heterogeneous image change detection. *Remote Sens. Environ.* 314, 114373.
- Zheng, Z., Ma, A., Zhang, L., Zhong, Y., 2021. Change is everywhere: Single-temporal supervised object change detection in remote sensing imagery. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 15193–15202.
- Zheng, Z., Zhong, Y., Wang, J., Ma, A., Zhang, L., 2023. FarSeg++: Foreground-Aware relation network for geospatial object segmentation in high spatial resolution remote sensing imagery. *IEEE Trans. Pattern Anal. Mach. Intell.* 45 (11), 13715–13729.
- Zheng, Z., Zhong, Y., Zhao, J., Ma, A., Zhang, L., 2024. Unifying remote sensing change detection via deep probabilistic change models: From principles, models to applications. *ISPRS J. Photogramm. Remote Sens.* 215, 239–255.
- Zhu, L., Sun, W., Fan, D., Xing, H., Liu, X., 2024. Unsupervised spatial self-similarity difference-based change detection method for multi-source heterogeneous images. *Pattern Recognit.* 149, 110237, [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0031320323009342>.