

Article

MSGATN: A Superpixel-Based Multi-Scale Siamese Graph Attention Network for Change Detection in Remote Sensing Images

Wenjing Shuai ^{1,*} , Fenlong Jiang ² , Hanhong Zheng ²  and Jianzhao Li ² 

¹ School of Electronic Engineering, Xidian University, Xi'an 710121, China

² Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education, School of Electronic Engineering, Xidian University, Xi'an 710071, China; jiangfenlong@outlook.com (F.J.); hanhong_zheng@163.com (H.Z.); 19jzli@stu.xidian.edu.cn (J.L.)

* Correspondence: wjshuai@xidian.edu.cn; Tel.: +86-182-2903-4065

Abstract: With the rapid development of Earth observation technology, how to effectively and efficiently detect changes in multi-temporal images has become an important but challenging problem. Relying on the advantages of high performance and robustness, object-based change detection (CD) has become increasingly popular. By analyzing the similarity of local pixels, object-based CD aggregates similar pixels into one object and takes it as the basic processing unit. However, object-based approaches often have difficulty capturing discriminative features, as irregular objects make processing difficult. To address this problem, in this paper, we propose a novel superpixel-based multi-scale Siamese graph attention network (MSGATN) which can process unstructured data natively and extract valuable features. First, a difference image (DI) is generated by Euclidean distance between bitemporal images. Second, superpixel segmentation is employed based on DI to divide each image into many homogeneous regions. Then, these superpixels are used to model the problem by graph theory to construct a series of nodes with the adjacency between them. Subsequently, the multi-scale neighborhood features of the nodes are extracted through applying a graph convolutional network and concatenated by an attention mechanism. Finally, the binary change map can be obtained by classifying each node by some fully connected layers. The novel features of MSGATN can be summarized as follows: (1) Training in multi-scale constructed graphs improves the recognition over changed land cover of varied sizes and shapes. (2) Spectral and spatial self-attention mechanisms are exploited for a better change detection performance. The experimental results on several real datasets show the effectiveness and superiority of the proposed method. In addition, compared to other recent methods, the proposed can demonstrate very high processing efficiency and greatly reduce the dependence on labeled training samples in a semisupervised training fashion.

Keywords: change detection; superpixel segmentation; graph attention network; remote sensing images



Citation: Shuai, W.; Jiang, F.; Zheng, H.; Li, J. MSGATN: A Superpixel-Based Multi-Scale Siamese Graph Attention Network for Change Detection in Remote Sensing Images. *Appl. Sci.* **2022**, *12*, 5158. <https://doi.org/10.3390/app12105158>

Academic Editor: Junseop Lee

Received: 23 April 2022

Accepted: 16 May 2022

Published: 20 May 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

With the continuous collection of massive multi-temporal remote sensing images, such as multi-spectral [1,2], synthetic aperture radar (SAR) [3], hyperspectral [4], and unmanned aerial vehicle (UAV) images [5], these multi-temporal remote sensing images have been promoted in practical applications. In this data context, change detection (CD) is one of the most meaningful technologies, which aims to quantitatively and qualitatively obtain the change information of ground objects by analyzing bitemporal remote sensing images. In many practical situations, these changes have potential significance, such as urban development planning, natural disaster assessment, dynamic monitoring of ecological environment, and natural resource management [6–8].

In the early stages, in order to obtain land cover change information, the traditional CD technology usually includes the following steps. First, the bitemporal image needs to

be preprocessed, including radiation correction, ensemble correction, and spatial registration [9,10]. Second, a difference image (DI) between bitemporal images can be acquired by image ratio [11], image difference [12], change vector analysis [13,14], etc. Finally, a threshold or a clustering algorithm is applied to segment DI into binary change map (BCM), such as Otsu [15,16], double-window flexible pace search [17,18], K-means [19], fuzzy c-means [20,21], and so on. Since bitemporal images are usually collected under different imaging conditions, such as illumination, season, etc., the different images may contain a large number of spurious differences [22]. Moreover, these methods usually use pixels as processing units, and either thresholding or clustering can cause a lot of noise in the results of these methods.

To address the limitations of pixel-level methods, many scholars have made great efforts in CD and proposed various object-based CD methods [23,24]. In general, object-based methods first need to segment the image to obtain multi-scale objects. Universal image segmentation techniques include fractal net evolution segmentation approach [25], simple linear iterative clustering (SLIC) superpixel segmentation [26], etc. These approaches frequently generate multi-scale objects or superpixels through region growing, i.e., objects or superpixels are obtained by gradual pixel binning with similar spectral values. Therefore, each superpixel or object is composed of a homogeneous set of pixels. The CD can then be achieved by the object for the image analysis and processing unit. For example, in the early stages, Jungcho et al. proposed an object-based CD based on correlation image analysis and image segmentation [27]. An object-based approach is based on multiple classifiers and multi-scale uncertainty analysis for CD with high-resolution (HR) remote sensing images [28]. Recently, some novel object-based approaches have made some efforts. For instance, Lv et al. promoted an object-oriented key point vector distance to obtain binary CD [29]. This method can significantly improve the performance of the difference image, as it measures the difference between the key-points vectors of two objects in bitemporal images. Similar methods are available in [30–32].

Although the aforementioned approaches have made remarkable progress, some limitations are still unavoidable. These limitations mainly include the following three aspects:

- Traditional methods are difficult to deal with and analyze irregular objects effectively as the multi-scale objects or the superpixels represent unstructured data. Therefore, there is still a lack of effective representative feature extraction approach for unstructured data.
- Image segmentation itself is a challenging task, and usually some parameters need to be adjusted to obtain better segmentation results. Moreover, the error of image segmentation may accumulate in the change detection task to some extent. Therefore, object-based change detection is severely limited by the performance of image segmentation.
- Object-based CD approaches generally require more complex frameworks. This results in a lower degree of automation of the entire CD framework due to the need to individually perform image segmentation algorithms and select appropriate segmentation parameters.

With the popularization of deep learning technology, the methods based on deep neural networks have been widely used in change detection. In particular, the graph neural networks (GNNs) have been noticed due to their excellent performance in unstructured data classification. Recently, GNNs have been successfully applied to image classification [33,34] and change detection [35,36]. Specifically, in [37], graph convolutional networks (GCNs) are utilized to extract the features of different types for hyperspectral image classification. Saha et al. proposed a semisupervised CD approach based on GCNs [38], which adopts multi-scale parcel segmentation to encode multi-temporal images as a graph. However, there are still few studies on GCN-based CD at present. Therefore, GCNs-based CD still needs continuous and further research.

Considering the excellent performance of GCNs in solving image classification, we are able to model the change detection task as a graph node classification task for improving the performance of CD. With this motivation, the paper proposes a novel multi-scale superpixel

graph attention network (MSGATN), which can process unstructured data natively and extract valuable features. In the proposed method, a difference image (DI) is firstly obtained by Euclidean distance between bitemporal images. Then, an SLIC algorithm is employed to divide the DI into many homogeneous superpixels. Subsequently, these superpixels are exploited to model the problem by graph theory to build a series of nodes with Based on this, the multi-scale features of each node are captured by a graph attention network (GATN). Finally, the binary change map (BCM) is generated by classifying each node using some fully connected layers.

The contributions of the proposed MSGATN approach are summarized as follows:

- (1) We propose a network model based on graph theory, which can process the unstructured data of objects with irregular boundaries in OBCD and consider the adjacency relationship between objects.
- (2) The proposed method is inductive, which can simultaneously adapt to graphs of different scales. Therefore, our proposed MSGATN can exploit the constructed graphs of various scales, thus improving the abilities of representation and generalization.
- (3) Experiments on several real datasets obtained from different sensors demonstrate that the proposed MSGATN has high efficiency and performance, as well as certain generalization.

The rest of this paper is organized as follows. Section 2 briefly introduces some related works. In Section 3, our method is described in detail. Section 4 provides the experimental settings and results. Finally, the conclusions and future works are given in Section 5.

2. Related Work

2.1. Deep-Learning-Based CD Methods

In recent years, deep learning technology has become a new favorite in the field of CD [39,40], especially convolutional neural networks (CNNs). These deep-learning-based methods can be roughly summarized into two categories, i.e., image-level methods and patch-level methods.

(1) Image-level methods: This category of method is to acquire semantic change information by analyzing a complete bitemporal image at a time [41]. Hence, that usually requires a large number of pairs of manually labeled training image pairs. For example, Ji et al. proposed a Siamese U-Net with shared weights to acquire a building change map in an end-to-end manner [42]. Liu et al. devised a local-global pyramid network for building CD in [43]. In [44], a spatial-temporal attention-based network based on self-attention mechanism was applied to mine deep robust features for large image-to-image CD datasets. Although these approaches can achieve competitive performance, they often not only require a large number of manually labeled paired images to train the network, but also cost more storage space and computational resources.

(2) Patch-level methods: Different from image-level methods, this type of method indicates using local pixel patches or superpixels as analysis units, and capturing feature representations through convolution or fully connection to achieve CD. In the early stages, Gong et al. proposed a novel CD method based on deep learning [45], which can avoid the effect of the DI to provide a better change detection performance. In [46], a Gabor-based PCANet (GaborPCANet) was promoted for CD in SAR images, which utilizes PCA filters as convolutional filters to capture the image features. A convolutional-wavelet neural network (CWNN) was devised to detect sea ice change detection from SAR images in [47]. Jiang et al. developed a semisupervised multiple CD approach, which can detect multiple changes using only a very limited samples by training a generative adversarial network [48]. This approach introduces dual-tree complex wavelet transform into CNNs to reduce the effect of the speckle noise, thus improving detection performance. However, these methods based on local pixel patches are still limited by the selection of regular windows. To alleviate this limitation, superpixels-based CD methods have received attention, which aim to use superpixels as analysis units to capture more representative features through CNNs. To achieve this, recent methods have made further efforts. For instance, Gong et al.

presented a superpixel-based difference representation learning to extract semantic change information between bitemporal images [49]. In [50], an end-to-end superpixel-enhanced CD network was designed, which combines an adaptive superpixel merging module to mine difference information for CD. Other methods refer to [51–53].

2.2. Graph Neural Networks

In the early stage, among the works concerning CD approach, graph theory-based approaches have been extensively used for CD [54,55]. For instance, in [56], a weighted graph was built to measure changes for CD with SAR images. Sun et al. proposed an iterative robust graph for unsupervised heterogeneous images CD [57]. This method constructs a robust K-nearest neighbor graph of bitemporal images, and calculates the difference image by comparing the graphs.

With the development of deep learning, a variant of CNNs, graph neural networks (GNNs), has received sustained attention in many applications [58–60]. In particular, graph convolutional networks (GCNs) have been successfully applied in the fields of remote sensing, such as remote sensing image retrieval [61], remote sensing image semantic segmentation [62], and hyperspectral image classification [63]. Specifically, GCNs are able to efficiently process graph-structured data by modeling the relationships between samples (or vertices). Therefore, GCNs can be naturally used to model remote spatial relationships in remote sensing images, which is not considered in CNNs. Recently, considering the previous GCNs-based research in the field of remote sensing, these methods have been developed and applied to CD tasks. For example, Wu et al. promoted a multi-scale GCN to detect land cover changes for CD in homogeneous and heterogeneous remote sensing images [64]. This approach constructs graph representations through object-wise high-level features generated by a pretrained U-Net. In [65], a multi-scale dynamic GCN was employed to mine the short-range and long-range contextual information. These GNNs-based methods have been initially applied to solve remote sensing image CD. However, there are still few CD methods for GNNs, and a large number of systematic theoretical studies and applications are still lacking. Therefore, further development of GNN-based CD methods has potential value.

3. Proposed Superpixel-Based MSGATN

3.1. Overview of the Proposed MSGATN

In this subsection, an overview of the proposed MSGATN is given briefly in Figure 1. Firstly, the difference intensity of bitemporal remote sensed images is obtained by Euclidean distance. Based on the pixel-wise similarity, the difference intensity map can be segmented to massive unstructured multi-scale superpixels of varied shapes and boundaries by simple linear iterative clustering (SLIC). With the segmented DI acquired, a region adjacency graph (RAG) can be constructed based on the mutual consistency of neighbor superpixels. The spatial-temporal relationships between these superpixels can be well modeled by the edges of constructed DI RAG. Then, the bitemporal remote sensing images are also segmented into superpixels with the guidance of the segmentation information extracted in DI superpixel segmentation. Several significant statistical characteristics, i.e., minimum, maximum, mean, standard deviation, skewness, and kurtosis can further represent the features of multi-scale bitemporal superpixels. As a result, the input graph of graph attention network (GATN) can be constructed by the nodes obtained from features of bitemporal superpixels and the edges acquired in the RAG of DI superpixels. Finally, superpixel-level prediction is obtained by GATN and remapped to form the pixel-level change map. The detailed inference process of MSGATN can be illustrated in Algorithm 1.

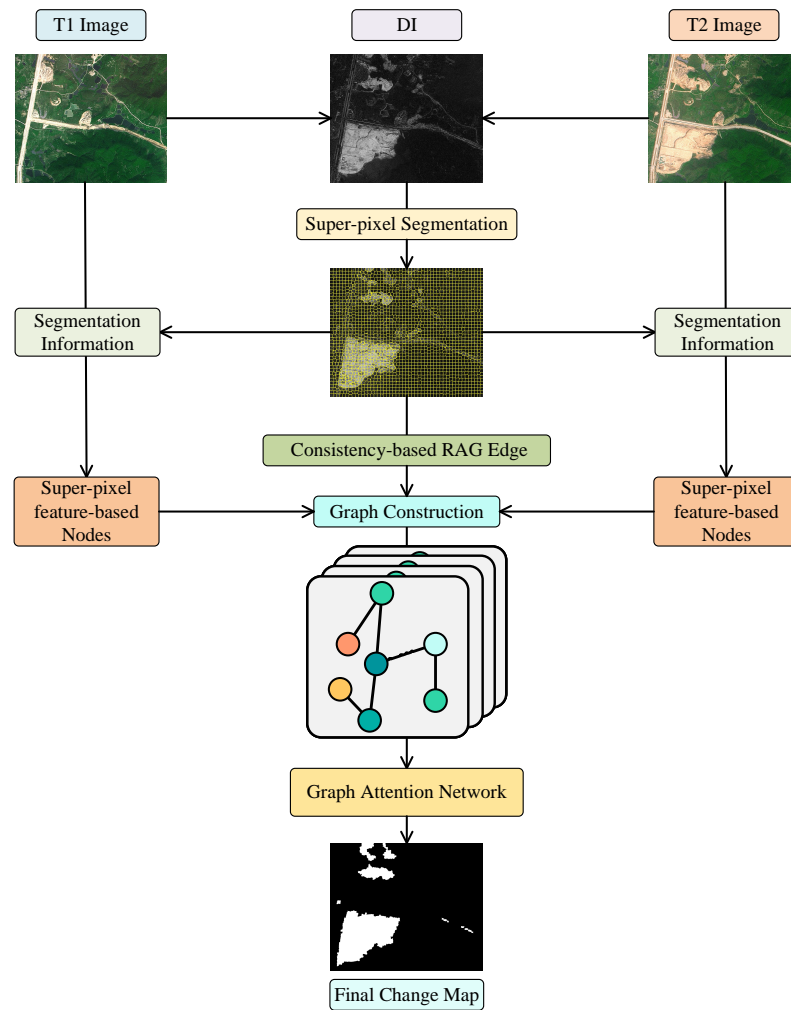


Figure 1. The framework and procedure of MSGATN.

Algorithm 1 Inference process of MSGATN

Input: T^1, T^2 : the bitemporal images.

- 1: Begin
- 2: $DI \leftarrow \text{Euclidean_Distance}(T1, T2)$; // obtain the difference intensity
- 3: $S^{DI} \leftarrow \text{SLIC}(DI)$; // conduct superpixel segmentation over DI
- 4: $\mathcal{G}_{sp}(\mathcal{V}, \mathcal{E}_{sp}) \leftarrow \text{RAG}(S^{DI})$; // acquire the region adjacency graph of S^{DI}
- 5: $S^{T1} \leftarrow \text{superpixel_segmentation}(T1)$; // segment $T1$ according to S^{DI}
- 6: $S^{T2} \leftarrow \text{superpixel_segmentation}(T2)$; // segment $T2$ according to S^{DI}
- 7: $F^1 \leftarrow \text{feature_analyse}(S^{T1})$; // represent the significant features of S^{T1}
- 8: $F^2 \leftarrow \text{feature_analyse}(S^{T2})$; // represent the significant features of S^{T2}
- 9: $\mathcal{V}_f \leftarrow \text{concatenate}(F^1, F^2)$; // collect the superpixel-level bitemporal features
- 10: $\mathcal{G}_{input} \leftarrow (\mathcal{V}_f, \mathcal{E}_{sp})$; // construct the input graph for GATN
- 11: $\mathcal{F}_{output} \leftarrow \text{GATN}(\mathcal{G}_{input})$; // obtain superpixel-wise change map
- 12: $CM \leftarrow \text{remap}(\mathcal{F}_{output})$; // remap the superpixel prediction to acquire final CM

Output: CM: binary change map.

As shown in the procedure above, the proposed MSGATN firstly obtains the multi-scale unstructured features of bitemporal remote sensed images through superpixel segmentation, which further promotes the fine-grained CD prediction. Then, the mutual relationships inside these bitemporal superpixels are well represented and modeled by GATN. Given the overall framework and inference process of the proposed MSGATN, the detailed information of the proposed graph construction mechanism can be given in the following section.

3.2. Graph Construction

To acquire credible prior information for GATN, a preliminary but representative graph construction is indispensable. In the proposed graph construction method, difference intensity and bitemporal remote sensing images are integrated to obtain comprehensive non-local change information, which advances the changed region detection in GATN. The overall graph construction can be further illustrated by the following steps. Initially, the pixel-wise difference intensity $DI \in \mathbb{R}^{H \times W}$ can be represented as

$$distance = \sqrt[2]{(T1 - T2)^2} \quad (1)$$

$$DI = \frac{distance - \min(distance)}{\max(distance) - \min(distance)} \quad (2)$$

After the difference intensity is acquired, the multi-scale superpixel segmentation over DI can be given as

$$S^{DI} = \{p_1, p_2, \dots, p_{N_seg}\} = \text{SLIC}_{N_seg}(DI) \quad (3)$$

where $\text{SLIC}_{N_seg}(\cdot)$ denotes the simple linear iterative clustering with different numbers of segmented superpixels, and N_seg represents the number of superpixels. Generally, the more superpixels, the smaller they are. In this case, we can obtain multi-scale superpixels for the multi-scale feature recognition of the proposed MSGATN. Then, the region adjacency co-relations \mathcal{G}_{sp} of S^{DI} can be modeled as

$$\mathcal{G}_{sp}(\mathcal{V}, \mathcal{E}_{sp}) = \text{RAG}(S^{DI}) \quad (4)$$

in which $\text{RAG}(\cdot)$ represents the region adjacency graph construction operation. The edges \mathcal{E}_{sp} are exploited to model the local and non-local relations inside bitemporal superpixels. To achieve this, the bitemporal superpixels need to be firstly acquired as $S^{T1} = \{p_1^1, p_2^1, \dots, p_{N_seg}^1\}$ and $S^{T2} = \{p_1^2, p_2^2, \dots, p_{N_seg}^2\}$. Then, the bitemporal features of these superpixels can be denoted as

$$F^1 = \text{concat}[\min(S^{T1}), \max(S^{T1}), \text{mean}(S^{T1}), \text{std}(S^{T1}), \text{skew}(S^{T1}), \text{kur}(S^{T1})] \quad (5)$$

$$F^2 = \text{concat}[\min(S^{T2}), \max(S^{T2}), \text{mean}(S^{T2}), \text{std}(S^{T2}), \text{skew}(S^{T2}), \text{kur}(S^{T2})] \quad (6)$$

where $\text{concat}(\cdot)$ indicates a feature-level integration, and $\min(\cdot)$, $\max(\cdot)$, $\text{mean}(\cdot)$, $\text{std}(\cdot)$, $\text{skew}(\cdot)$, and $\text{kur}(\cdot)$ represent the superpixel-wise minimum, maximum, mean, standard deviation, skewness, and kurtosis, respectively. Given these dependable and discriminative features of bitemporal superpixels, the nodes can be obtained as follows:

$$\mathcal{V}_f = \text{concat}(F^1, F^2) \quad (7)$$

at which $\text{concat}(\cdot)$ denotes a feature-wise concatenation. With the nodes and edges obtained, the input multi-scale graphs for MSGATN can be constructed as

$$\mathcal{G}_{input}^{N_seg} = (\mathcal{V}_f, \mathcal{E}_{sp}) \quad (8)$$

With different N_{seg} , input graphs of varied scales can be provided for the proposed MSGATN; thus, the multi-scale change objects obtain a finer cognition.

3.3. Multi-Scale Siamese Graph Attention Network

In the proposed MSGATN, a GATN is employed to better reveal the relationships between multi-scale unstructured superpixels, and the graph attention mechanism is the linchpin of GATN. To further facilitate the understanding for the proposed MSGATN, the mathematical style of graph attention is given as follows: Let $f_i \in \mathbb{R}^{C_I}$ and $f_j \in \mathbb{R}^{C_I}$ be the feature vectors of current node i and its neighbor node j , respectively. Then, the edge score e_{ij} can be obtained by

$$e_{ij} = (\text{concat}(f_i W, f_j W)) A \quad (9)$$

where $W \in \mathbb{R}^{C_I \times C_O}$ and $A \in \mathbb{R}^{2C_O \times 1}$ are the learnable supervised parameters, $\text{concat}(\cdot)$ represents a feature-wise integration, and C_I and C_O denote the input and output feature lengths, respectively. With each e_{ij} acquired, the attention score a_{ij} can be given as follows:

$$a_{ij} = \frac{\exp(\text{LeakyReLU}(e_{ij}))}{\sum_k \exp(\text{LeakyReLU}(e_{ik}))} \quad (10)$$

where $\text{LeakyReLU}(\cdot)$ represents a nonlinear activation, and k denotes all the neighbor nodes of i . In the proposed MSGATN, the graph attention mechanism is widely used to refine the graph feature representation. To improve the recognition for multi-scale objects, the proposed network is trained over multi-scale graphs from superpixel segmentation of different superpixel numbers. In our method, the N_{seg} is set to 2000, 4000, and 6000 to obtain input graphs of different scales. Basically, GATN can tackle inductive tasks with graphs of varied scales. Based on this fact, the proposed MSGATN can learn multi-scale feature representation through training over several multi-scale constructed graphs in a Siamese framework.

4. Experiments

4.1. Dataset Descriptions

To further test and verify the ability of the proposed method, two extensively used remote sensing CD multi-spectral datasets, i.e., the Guangzhou city dataset and the Hongqi canal dataset, are exploited, which are shown in Figures 2 and 3.

4.1.1. Guangzhou City Dataset

This dataset is composed of a bitemporal multispectral image pair with the spatial resolution of 2.5 m, captured by the Systeme Probatoire d'Observation de la Terre 5 (SPOT-5) satellite. It depicts the land cover change over urban areas of Guangzhou City between October 2006 and October 2007, respectively, as shown in Figure 2. The bitemporal images are the size of 877×738 pixels, including red, green, and near-infrared bands. Its annotation focuses on vegetation change.

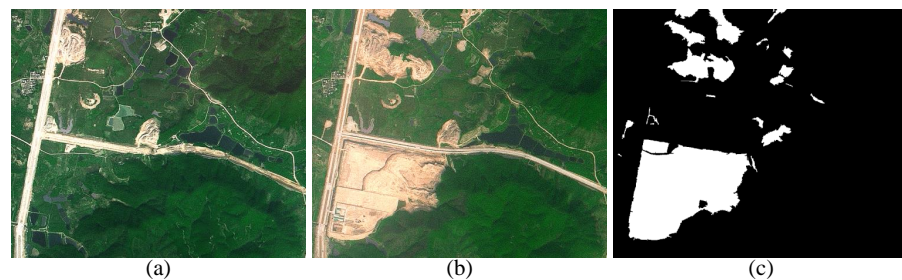


Figure 2. Guangzhou City dataset: (a) T_1 -time image, (b) T_2 -time image, (c) ground truth image.

4.1.2. Hongqi Canal Dataset

The second dataset, Hongqi Canal dataset, contains two high-resolution multispectral remote sensed images which focus on the region of Yellow River Estuary near the city of Dongying in China, as shown in Figure 3. The bitemporal images, which have the spatial resolution of 2m and the size of 539×543 , were acquired by GF-1 satellite on 9 December 2013 and 16 October 2015, respectively. It mainly describes the river changes of the Hongqi Canal settled in Xijiu village.

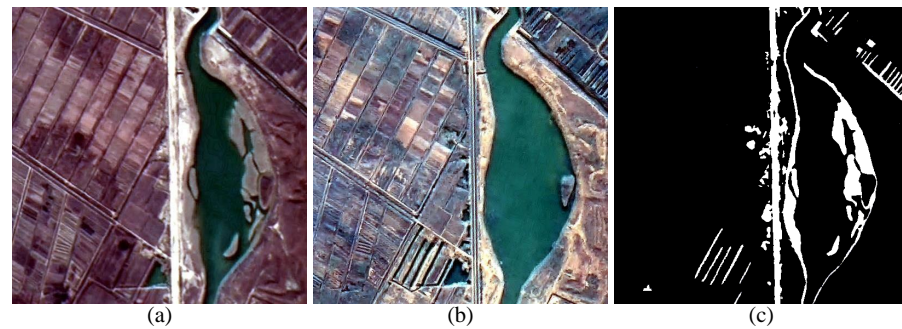


Figure 3. Hongqi Canal dataset: (a) T_1 -time image, (b) T_2 -time image, (c) ground truth image.

4.2. Comparative Methods and Related Settings

In the experiments, to evaluate the performance of the proposed MSGATN, we selected five related CD approaches for comparison with our MSGATN. All methods are described as follows:

- (1) PCA_K-means [19]: This approach is one of the popular unsupervised CD methods, which adopts principal component analysis (PCA) and k-means clustering to acquire binary change map. In this method, two parameters (h and s) should be set. For the Guangzhou City dataset, h and s are set to 9 and 3, respectively. For the Hongqi Canal dataset, h and s are set to 5 and 3, respectively.
- (2) ASEA [66]: It is a state-of-the-art method that exploits the contextual information around a pixel to improve detection accuracy. This method requires no parameter setting.
- (3) GaborPCANet [46]: This was proposed in [46]. It utilizes PCA filters as convolution kernels to obtain representative neighborhood features. In this approach, a parameter, patch size, is set to 5 for both experimental datasets.
- (4) DBN [49]: This is a superpixel-based method, which can acquire a better detection result by difference representation learning. For our experimental datasets, the parameter patch size is fixed to 5 in this method.
- (5) CWNN [47]: It devises a convolutional-wavelet neural network in SAR images. In the experiments, the parameter patch size is fixed to 7 for our datasets.
- (6) Proposed MSGATN: In our MSGATN, the number of superpixels is a hyperparameter. Specifically, in our method, we selected six scales of superpixel segmentation, which, respectively, include 1000, 2000, 3000, 4000, 5000, and 6000 superpixels, to train our MSGATN in a Siamese manner. For both experimental results, the results of 6000 superpixels are chosen to be compared with other methods.

4.3. Evaluation Criteria

To further evaluate the performance for CD, several widely used evaluation metrics, which are precision, recall, $F1$, and overall accuracy (OA), are employed. Their definitions are given as follows:

$$Precision = \frac{TP}{TP + FP} \quad (11)$$

$$Recall = \frac{TP}{TP + FN} \quad (12)$$

$$F1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (13)$$

$$OA = \frac{TP + TN}{TP + TN + FP + FN} \quad (14)$$

where TP , TN , FP , and FN are the numbers of true positive, true negative, false positive, false negative pixels, respectively. Based on these well-acknowledged evaluation metrics, the performance of different CD methods can be better revealed.

4.4. Comparative Results

In this subsection, the comparative results on two widely used CD datasets are given in detail. To further illustrate the proposed method, corresponding analysis will be given in detail. Detailed visual and quantized results and analysis are presented as follows.

4.4.1. Results on the Guangzhou City Dataset

The visualized and quantitative comparison results over the Guangzhou City dataset are shown in Figure 4 and Table 1, respectively. From the quantitative comparison, our MSGATN can achieve the best $F1$ and OA (90.54% and 97.19%). However, DBN and GaborPCANet achieved the best precision and recall, respectively. Although the proposed MSGATN does not achieve the best precision and recall, our method still provides relatively reliable performance in terms of precision and recall. For example, compared with the DBN, despite DBN reaching the best precision (98.05%), it obtained the second-worst performance in recall (78.51%). Therefore, our MSGATN can acquire more balanced performance for the four metrics. Different from other approaches, the proposed MSGATN adopts a multi-scale graph attention network to effectively capture the representative features of unstructured data, thereby improving the detection accuracy. Regarding the visual results, the proposed MSGATN exhibits the fewest false detections compared to the other five methods. Specifically, the GaborPCANet presents many false alarms compared to the proposed MSGATN. Moreover, although the visual results of the DBN show fewer false alarms, a large number of missed pixels are unavoidable. Compared with these methods, our proposed MSGATN can obtain a more balanced performance in terms of false detections and missed detections. Moreover, the proposed MSGATN can provide more complete change information compared with other methods, except GaborPCANet. Overall, the visual results also yielded similar conclusions to the quantitative comparisons.

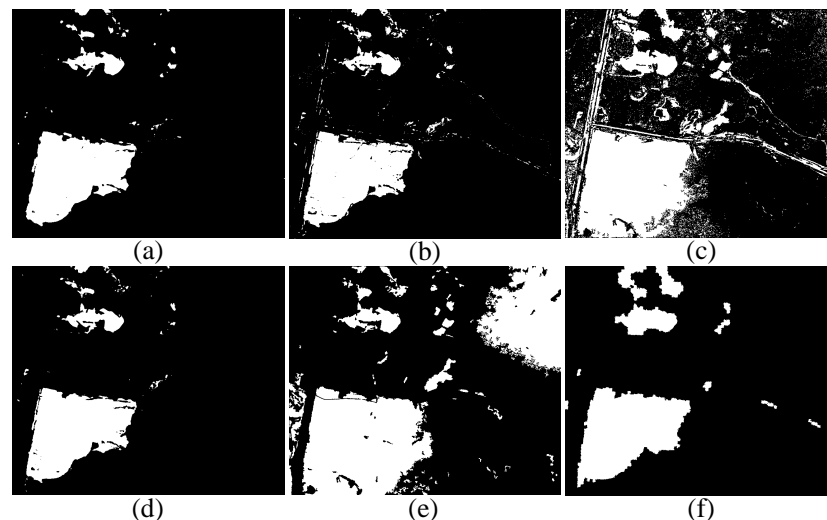


Figure 4. The results of different methods on the Guangzhou City dataset: (a) PCA_K-means, (b) ASEA, (c) GaborPCANet, (d) DBN, (e) CWNN, and (f) proposed MSGATN.

Table 1. Accuracy comparison (in %) of different methods on the Guangzhou City dataset. The best evaluation value are presented in bold for different metrics.

Methods	Precision	Recall	F1	OA
PCA_K-means	97.33	78.22	86.74	96.43
ASEA	95.45	79.73	86.88	96.40
GaborPCANet	51.61	94.98	66.88	85.94
DBN	98.05	78.51	87.20	96.55
CWNN	42.19	87.17	56.86	80.23
Proposed MSGATN	91.04	90.04	90.54	97.19

4.4.2. Results on the Hongqi Canal Dataset

As presented in Table 2, the proposed MSGATN achieve the overall superiority over the Hongqi Canal dataset compared to other selected CD methods. That is, our method outperforms the other methods in all evaluation indicators, i.e., precision, recall, F1, and OA, with a great gap. More exactly, the best precision (80.96%) and recall (57.17%) are achieved by the proposed MSGATN, which leads to the best F1 (67.02%) for our method. It indicates that the proposed MSGATN can capture finer complete land cover and acquire better mapping for changed multi-scale objects with the help of input graphs with varied scales, and similar conclusions can be discovered in the visualized CD results depicted in Figure 5. Given the fact that the annotation of the Hongqi Canal dataset mainly focuses on the river change, massive false alarms can be found in the CMs generated by other methods. These false alarms are basically caused by the unchanged farmland around the canal. However, they are well filtered out in the proposed MSGATN, which can be attributed to the finer feature representation of our method. As a result, the river course change in the Hongqi Canal dataset is well denoted by the proposed MSGATN, which suggests the advantage of our method.

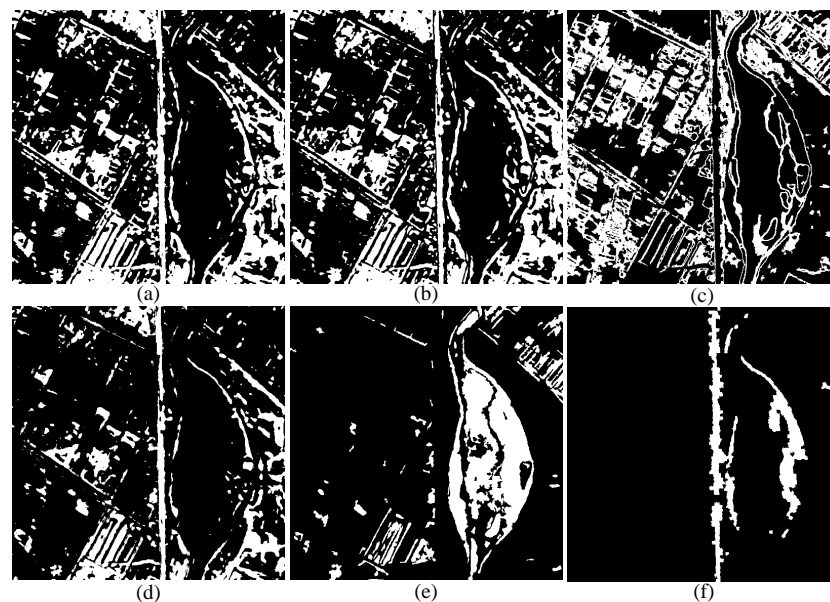


Figure 5. The results of different methods on the Hongqi Canal dataset: (a) PCA_K-means, (b) ASEA, (c) GaborPCANet, (d) DBN, (e) CWNN, and (f) proposed MSGATN.

Table 2. Accuracy comparison (in %) of different methods on the Hongqi Canal dataset. The best evaluation value are presented in bold for different metrics.

Methods	Precision	Recall	F1	OA
PCA_K-means	15.67	47.28	23.53	71.57
ASEA	16.26	50.69	24.62	71.28
GaborPCANet	3.71	12.89	5.76	60.99
DBN	19.17	34.53	24.66	80.47
CWNN	31.91	56.12	40.68	84.86
Proposed MSGATN	80.96	57.17	67.02	94.80

4.5. Parameters Analysis of the Proposed MSGATN on the Guangzhou Dataset

To further investigate the effectiveness of the proposed MSGATN, parameters analyses are performed on the Guangzhou dataset in this section. In our MSGATN, the number of superpixels is a hyperparameter. Furthermore, we selected six scales of superpixel segmentation, which, respectively, include 1000, 2000, 3000, 4000, 5000, and 6000 superpixels (as shown in Figure 6), to train our MSGATN in a Siamese manner. Generally, a larger number of superpixels indicates a smaller segmentation scale. Conversely, a smaller number of superpixels indicates a larger segmentation scale. Thanks to the characteristics of the inductive GATN, our MSGATN can easily exploit multi-scale superpixel features. By this way, features of different scales can be considered in our method. In this context, different BCMs can be generated by the proposed MSGATN at each scale, as presented in Figure 6. As the number of superpixels increases, the scale of superpixels also becomes finer. Similarly, the BCM of each scale is also finer as the number of superpixels enlarges for our MSGATN.

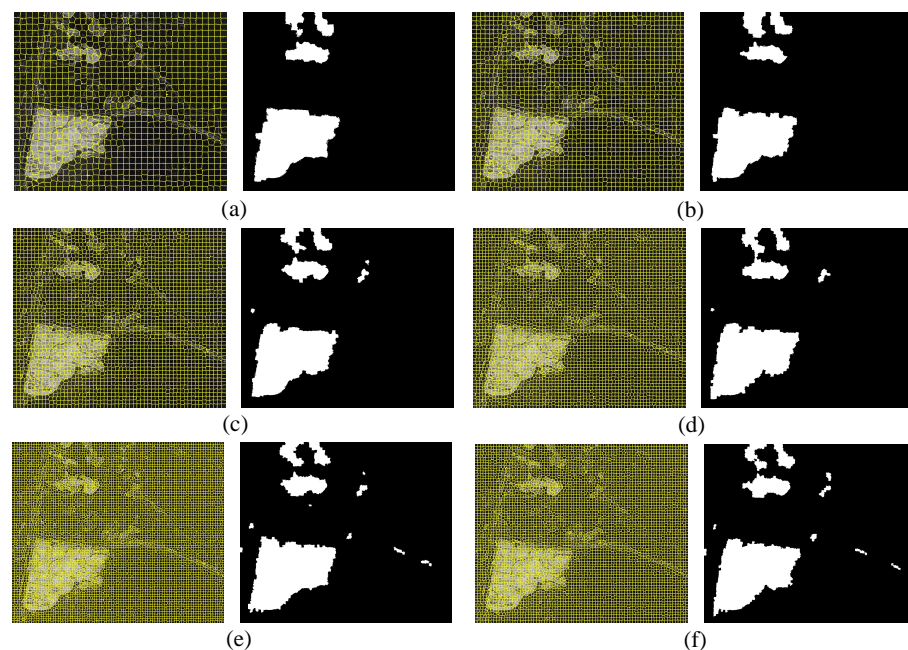


Figure 6. Segmentation results and the corresponding change detection results of different superpixel numbers in the proposed MSGATN on the Guangzhou dataset: (a) 1000 superpixels, (b) 2000 superpixels, (c) 3000 superpixels, (d) 4000 superpixels, (e) 5000 superpixels, (f) 6000 superpixels.

Figure 7 more intuitively demonstrates the relationship between the number of superpixels and detection accuracy. Concretely, as the number of superpixels increases, all metrics show an upward trend. However, if the number of superpixels exceeds 3000, the accuracy gradually decreases. Hence, the performance of the proposed MSGATN may not continue to increase as the number of superpixels increases. Moreover, more superpixels can lead to larger graph structures, which can significantly increase the computational

cost. According to the above analysis, the number of superpixels cannot be continuously increased in our method.

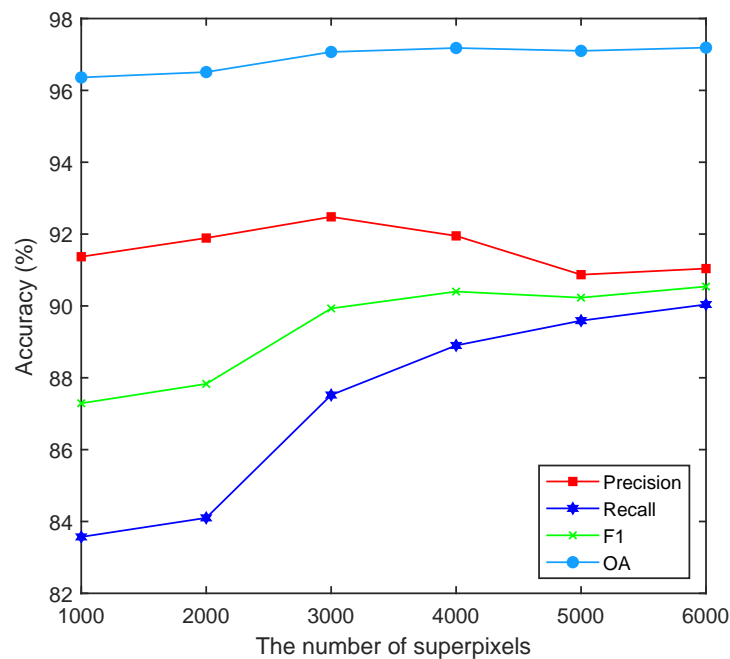


Figure 7. Relationship between change detection accuracy and superpixels numbers for the proposed MSGATN on the Guangzhou dataset.

5. Conclusions

In this work, a novel superpixel-based multi-scale Siamese graph attention network (MSGATN) is proposed for change detection in high-resolution remote sensed imagery. In the proposed method, superpixel segmentation is exploited to aggregate homogeneous difference information to construct heterogeneous change information for a better recognition of multi-scale changed land cover. In addition, multi-scale superpixel-constructed graphs are introduced to a graph attention network (GATN) in a Siamese framework, which further facilitates the cognition of multi-scale objects for the GATN, thus improving the performance. The proposed MSGATN is validated over two widely used change detection datasets, and compared to several comparative change detection methods. Corresponding results indicate that the proposed method outperforms other methods over all selected evaluation metrics.

In the future work, efforts can be made to achieve a more fine-grained changed land cover annotation in an unsupervised framework, which can be less time-consuming and laboring in practical applications.

Author Contributions: Conceptualization, W.S. and F.J.; methodology, F.J. and H.Z.; validation, H.Z.; investigation, J.L.; writing—original draft preparation, W.S.; writing—review and editing, W.S. and F.J. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Natural Science Foundation of China (61906148).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Lv, Z.; Li, G.; Yan, J.; Benediktsson, J.A.; You, Z. Training Samples Enriching Approach for Classification Improvement of VHR Remote Sensing Image. *IEEE Geosci. Remote Sens. Lett.* **2021**, *19*, 1–5. [\[CrossRef\]](#)
2. Wu, Y.; Xiao, Z.; Liu, S.; Miao, Q.; Ma, W.; Gong, M.; Xie, F.; Zhang, Y. A Two-Step Method for Remote Sensing Images Registration Based on Local and Global Constraints. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 5194–5206. [\[CrossRef\]](#)

3. Wu, Y.; Ma, W.; Gong, M.; Su, L.; Jiao, L. A novel point-matching algorithm based on fast sample consensus for image registration. *IEEE Geosci. Remote Sens. Lett.* **2014**, *12*, 43–47. [\[CrossRef\]](#)
4. Lv, Z.; Li, G.; Jin, Z.; Benediktsson, J.A.; Foody, G.M. Iterative training sample expansion to increase and balance the accuracy of land classification from VHR imagery. *IEEE Trans. Geosci. Remote Sens.* **2020**, *59*, 139–150. [\[CrossRef\]](#)
5. Liu, T.; Gong, M.; Jiang, F.; Zhang, Y.; Li, H. Landslide Inventory Mapping Method Based on Adaptive Histogram-Mean Distance With Bitemporal VHR Aerial Images. *IEEE Geosci. Remote Sens. Lett.* **2021**, *19*, 1–5. [\[CrossRef\]](#)
6. Zhu, Z. Change detection using landsat time series: A review of frequencies, preprocessing, algorithms, and applications. *ISPRS J. Photogramm. Remote Sens.* **2017**, *130*, 370–384. [\[CrossRef\]](#)
7. Shi, W.; Zhang, M.; Zhang, R.; Chen, S.; Zhan, Z. Change detection based on artificial intelligence: State-of-the-art and challenges. *Remote Sens.* **2020**, *12*, 1688. [\[CrossRef\]](#)
8. Tewkesbury, A.P.; Comber, A.J.; Tate, N.J.; Lamb, A.; Fisher, P.F. A critical synthesis of remotely sensed optical image change detection techniques. *Remote Sens. Environ.* **2015**, *160*, 1–14. [\[CrossRef\]](#)
9. Wu, Y.; Miao, Q.; Ma, W.; Gong, M.; Wang, S. PSOSAC: Particle swarm optimization sample consensus algorithm for remote sensing image registration. *IEEE Geosci. Remote Sens. Lett.* **2017**, *15*, 242–246. [\[CrossRef\]](#)
10. Wu, Y.; Liu, J.W.; Zhu, C.Z.; Bai, Z.F.; Miao, Q.G.; Ma, W.P.; Gong, M.G. Computational intelligence in remote sensing image registration: A survey. *Int. J. Autom. Comput.* **2021**, *18*, 1–17. [\[CrossRef\]](#)
11. Gong, M.; Cao, Y.; Wu, Q. A neighborhood-based ratio approach for change detection in SAR images. *IEEE Geosci. Remote Sens. Lett.* **2011**, *9*, 307–311. [\[CrossRef\]](#)
12. Lv, Z.; Liu, T.; Zhang, P.; Atli Benediktsson, J.; Chen, Y. Land cover change detection based on adaptive contextual information using bitemporal remote sensing images. *Remote Sens.* **2018**, *10*, 901. [\[CrossRef\]](#)
13. Liu, S.; Bruzzone, L.; Bovolo, F.; Zanetti, M.; Du, P. Sequential spectral change vector analysis for iteratively discovering and detecting multiple changes in hyperspectral images. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 4363–4378. [\[CrossRef\]](#)
14. Thonfeld, F.; Feilhauer, H.; Braun, M.; Menz, G. Robust Change Vector Analysis (RCVA) for multi-sensor very high resolution optical satellite data. *Int. J. Appl. Earth Obs. Geoinf.* **2016**, *50*, 131–140. [\[CrossRef\]](#)
15. Otsu, N. A threshold selection method from gray-level histograms. *IEEE Trans. Syst. Man Cybern.* **1979**, *9*, 62–66. [\[CrossRef\]](#)
16. Lv, Z.; Liu, T.; Atli Benediktsson, J.; Lei, T.; Wan, Y. Multi-scale object histogram distance for LCCD using bitemporal very-high-resolution remote sensing images. *Remote Sens.* **2018**, *10*, 1809. [\[CrossRef\]](#)
17. Lv, Z.Y.; Liu, T.F.; Zhang, P.; Benediktsson, J.A.; Lei, T.; Zhang, X. Novel adaptive histogram trend similarity approach for land cover change detection by using bitemporal very-high-resolution remote sensing images. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 9554–9574. [\[CrossRef\]](#)
18. Lv, Z.; Liu, T.; Shi, C.; Benediktsson, J.A. Local histogram-based analysis for detecting land cover change using VHR remote sensing images. *IEEE Geosci. Remote Sens. Lett.* **2020**, *18*, 1284–1287. [\[CrossRef\]](#)
19. Celik, T. Unsupervised change detection in satellite images using principal component analysis and *k*-means clustering. *IEEE Geosci. Remote Sens. Lett.* **2009**, *6*, 772–776. [\[CrossRef\]](#)
20. Ghosh, A.; Mishra, N.S.; Ghosh, S. Fuzzy clustering algorithms for unsupervised change detection in remote sensing images. *Inf. Sci.* **2011**, *181*, 699–715. [\[CrossRef\]](#)
21. Gong, M.; Zhou, Z.; Ma, J. Change detection in synthetic aperture radar images based on image fusion and fuzzy clustering. *IEEE Trans. Image Process.* **2011**, *21*, 2141–2151. [\[CrossRef\]](#) [\[PubMed\]](#)
22. Wu, Y.; Li, J.; Yuan, Y.; Qin, A.; Miao, Q.G.; Gong, M.G. Commonality autoencoder: Learning common features for change detection from heterogeneous images. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**. [\[CrossRef\]](#) [\[PubMed\]](#)
23. Chen, G.; Hay, G.J.; Carvalho, L.M.; Wulder, M.A. Object-based change detection. *Int. J. Remote Sens.* **2012**, *33*, 4434–4457. [\[CrossRef\]](#)
24. Hussain, M.; Chen, D.; Cheng, A.; Wei, H.; Stanley, D. Change detection from remotely sensed images: From pixel-based to object-based approaches. *ISPRS J. Photogramm. Remote Sens.* **2013**, *80*, 91–106. [\[CrossRef\]](#)
25. Chen, Q.; Li, L.; Xu, Q.; Yang, S.; Shi, X.; Liu, X. Multi-feature segmentation for high-resolution polarimetric SAR data based on fractal net evolution approach. *Remote Sens.* **2017**, *9*, 570. [\[CrossRef\]](#)
26. Achanta, R.; Shaji, A.; Smith, K.; Lucchi, A.; Fua, P.; Süsstrunk, S. SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE Trans. Pattern Anal. Mach. Intell.* **2012**, *34*, 2274–2282. [\[CrossRef\]](#)
27. Im, J.; Jensen, J.; Tullis, J. Object-based change detection using correlation image analysis and image segmentation. *Int. J. Remote Sens.* **2008**, *29*, 399–423. [\[CrossRef\]](#)
28. Tan, K.; Zhang, Y.; Wang, X.; Chen, Y. Object-based change detection using multiple classifiers and multi-scale uncertainty analysis. *Remote Sens.* **2019**, *11*, 359. [\[CrossRef\]](#)
29. Lv, Z.; Liu, T.; Benediktsson, J.A. Object-oriented key point vector distance for binary land cover change detection using VHR remote sensing images. *IEEE Trans. Geosci. Remote Sens.* **2020**, *58*, 6524–6533. [\[CrossRef\]](#)
30. Hao, M.; Zhou, M.; Jin, J.; Shi, W. An advanced superpixel-based Markov random field model for unsupervised change detection. *IEEE Geosci. Remote Sens. Lett.* **2019**, *17*, 1401–1405. [\[CrossRef\]](#)
31. Zhu, L.; Zhang, J.; Sun, Y. Remote Sensing Image Change Detection Using Superpixel Cosegmentation. *Information* **2021**, *12*, 94. [\[CrossRef\]](#)

32. Pang, S.; Hu, X.; Zhang, M.; Cai, Z.; Liu, F. Co-segmentation and superpixel-based graph cuts for building change detection from bitemporal digital surface models and aerial images. *Remote Sens.* **2019**, *11*, 729. [\[CrossRef\]](#)
33. Cai, W.; Wei, Z. Remote sensing image classification based on a cross-attention mechanism and graph convolution. *IEEE Geosci. Remote Sens. Lett.* **2020**, *19*, 1–5. [\[CrossRef\]](#)
34. Du, X.; Zheng, X.; Lu, X.; Doudkin, A.A. Multisource remote sensing data classification with graph fusion network. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 10062–10072. [\[CrossRef\]](#)
35. Wang, R.; Wang, L.; Dong, P.; Jiao, L.; Chen, J.W. Graph-Level Neural Network for SAR Image Change Detection. In Proceedings of the 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, Brussels, Belgium, 11–16 July 2021; pp. 3785–3788.
36. Kalinicheva, E.; Ienco, D.; Sublime, J.; Trocan, M. Unsupervised change detection analysis in satellite image time series using deep learning combined with graph-based approaches. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *13*, 1450–1466. [\[CrossRef\]](#)
37. Hong, D.; Gao, L.; Yao, J.; Zhang, B.; Plaza, A.; Chanussot, J. Graph convolutional networks for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* **2020**, *59*, 5966–5978. [\[CrossRef\]](#)
38. Saha, S.; Mou, L.; Zhu, X.X.; Bovolo, F.; Bruzzone, L. Semisupervised change detection using graph convolutional network. *IEEE Geosci. Remote Sens. Lett.* **2020**, *18*, 607–611. [\[CrossRef\]](#)
39. Liu, S.; Marinelli, D.; Bruzzone, L.; Bovolo, F. A review of change detection in multitemporal hyperspectral images: Current techniques, applications, and challenges. *IEEE Geosci. Remote Sens. Mag.* **2019**, *7*, 140–158. [\[CrossRef\]](#)
40. ZhiYong, L.; Liu, T.; Benediktsson, J.A.; Falco, N. Land cover change detection techniques: Very-high-resolution optical images: A review. *IEEE Geosci. Remote Sens. Mag.* **2021**, *10*, 44–63.
41. Zheng, H.; Gong, M.; Liu, T.; Jiang, F.; Zhan, T.; Lu, D.; Zhang, M. HFA-Net: High Frequency Attention Siamese Network for Building Change Detection in VHR Remote Sensing Images. *Pattern Recognit.* **2022**, *129*, 108717. [\[CrossRef\]](#)
42. Ji, S.; Wei, S.; Lu, M. Fully convolutional networks for multisource building extraction from an open aerial and satellite imagery data set. *IEEE Trans. Geosci. Remote Sens.* **2018**, *57*, 574–586. [\[CrossRef\]](#)
43. Liu, T.; Gong, M.; Lu, D.; Zhang, Q.; Zheng, H.; Jiang, F.; Zhang, M. Building change detection for vhr remote sensing images via local-global pyramid network and cross-task transfer learning strategy. *IEEE Trans. Geosci. Remote Sens.* **2021**, *60*, 1–17. [\[CrossRef\]](#)
44. Chen, H.; Shi, Z. A spatial-temporal attention-based method and a new dataset for remote sensing image change detection. *Remote Sens.* **2020**, *12*, 1662. [\[CrossRef\]](#)
45. Gong, M.; Zhao, J.; Liu, J.; Miao, Q.; Jiao, L. Change detection in synthetic aperture radar images based on deep neural networks. *IEEE Trans. Neural Netw. Learn. Syst.* **2015**, *27*, 125–138. [\[CrossRef\]](#)
46. Gao, F.; Dong, J.; Li, B.; Xu, Q. Automatic change detection in synthetic aperture radar images based on PCANet. *IEEE Geosci. Remote Sens. Lett.* **2016**, *13*, 1792–1796. [\[CrossRef\]](#)
47. Gao, F.; Wang, X.; Gao, Y.; Dong, J.; Wang, S. Sea ice change detection in SAR images based on convolutional-wavelet neural networks. *IEEE Geosci. Remote Sens. Lett.* **2019**, *16*, 1240–1244. [\[CrossRef\]](#)
48. Jiang, F.; Gong, M.; Zhan, T.; Fan, X. A semisupervised GAN-based multiple change detection framework in multi-spectral images. *IEEE Geosci. Remote Sens. Lett.* **2019**, *17*, 1223–1227. [\[CrossRef\]](#)
49. Gong, M.; Zhan, T.; Zhang, P.; Miao, Q. Superpixel-based difference representation learning for change detection in multispectral remote sensing images. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 2658–2673. [\[CrossRef\]](#)
50. Zhang, H.; Lin, M.; Yang, G.; Zhang, L. ESCNet: An End-to-End Superpixel-Enhanced Change Detection Network for Very-High-Resolution Remote Sensing Images. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**, 1–15. [\[CrossRef\]](#)
51. Gong, M.; Jiang, F.; Qin, A.; Liu, T.; Zhan, T.; Lu, D.; Zheng, H.; Zhang, M. A Spectral and Spatial Attention Network for Change Detection in Hyperspectral Images. *IEEE Trans. Geosci. Remote Sens.* **2021**, *60*, 1–15. [\[CrossRef\]](#)
52. Saha, S.; Bovolo, F.; Bruzzone, L. Unsupervised deep change vector analysis for multiple-change detection in VHR images. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 3677–3693. [\[CrossRef\]](#)
53. Du, B.; Ru, L.; Wu, C.; Zhang, L. Unsupervised deep slow feature analysis for change detection in multi-temporal remote sensing images. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 9976–9992. [\[CrossRef\]](#)
54. Vakalopoulou, M.; Karantzalos, K.; Komodakis, N.; Paragios, N. Graph-based registration, change detection, and classification in very high resolution multitemporal remote sensing data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2016**, *9*, 2940–2951. [\[CrossRef\]](#)
55. Wang, T.; Lu, G.; Liu, J.; Yan, P. Graph-based change detection for condition monitoring of rotating machines: Techniques for graph similarity. *IEEE Trans. Reliab.* **2018**, *68*, 1034–1049. [\[CrossRef\]](#)
56. Pham, M.T.; Mercier, G.; Michel, J. Change detection between SAR images using a pointwise approach and graph theory. *IEEE Trans. Geosci. Remote Sens.* **2015**, *54*, 2020–2032. [\[CrossRef\]](#)
57. Sun, Y.; Lei, L.; Guan, D.; Kuang, G. Iterative Robust Graph for Unsupervised Change Detection of Heterogeneous Remote Sensing Images. *IEEE Trans. Image Process.* **2021**, *30*, 6277–6291. [\[CrossRef\]](#)
58. Fan, X.; Gong, M.; Wu, Y.; Qin, A.; Xie, Y. Propagation Enhanced Neural Message Passing for Graph Representation Learning. *IEEE Trans. Knowl. Data Eng.* **2021**. [\[CrossRef\]](#)
59. Gong, M.; Zhou, H.; Qin, A.; Liu, W.; Zhao, Z. Self-Paced Co-Training of Graph Neural Networks for Semi-Supervised Node Classification. *IEEE Trans. Neural Netw. Learn. Syst.* **2022**. [\[CrossRef\]](#) [\[PubMed\]](#)
60. Fan, X.; Gong, M.; Xie, Y.; Jiang, F.; Li, H. Structured self-attention architecture for graph-level representation learning. *Pattern Recognit.* **2020**, *100*, 107084. [\[CrossRef\]](#)

61. Chaudhuri, U.; Banerjee, B.; Bhattacharya, A. Siamese graph convolutional network for content based remote sensing image retrieval. *Comput. Vis. Image Underst.* **2019**, *184*, 22–30. [[CrossRef](#)]
62. Ouyang, S.; Li, Y. Combining deep semantic segmentation network and graph convolutional neural network for semantic segmentation of remote sensing imagery. *Remote Sens.* **2020**, *13*, 119. [[CrossRef](#)]
63. Qin, A.; Shang, Z.; Tian, J.; Wang, Y.; Zhang, T.; Tang, Y.Y. Spectral–spatial graph convolutional networks for semisupervised hyperspectral image classification. *IEEE Geosci. Remote Sens. Lett.* **2018**, *16*, 241–245. [[CrossRef](#)]
64. Wu, J.; Li, B.; Qin, Y.; Ni, W.; Zhang, H.; Fu, R.; Sun, Y. A multiscale graph convolutional network for change detection in homogeneous and heterogeneous remote sensing images. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *105*, 102615. [[CrossRef](#)]
65. Tang, X.; Zhang, H.; Mou, L.; Liu, F.; Zhang, X.; Zhu, X.X.; Jiao, L. An Unsupervised Remote Sensing Change Detection Method Based on Multiscale Graph Convolutional Network and Metric Learning. *IEEE Trans. Geosci. Remote Sens.* **2021**, *60*, 1–15. [[CrossRef](#)]
66. Lv, Z.; Wang, F.; Liu, T.; Kong, X.; Benediktsson, J.A. Novel Automatic Approach for Land Cover Change Detection by Using VHR Remote Sensing Images. *IEEE Geosci. Remote Sens. Lett.* **2021**, *19*, 1–5. [[CrossRef](#)]