

MSGFNet for Remote Sensing Image Change Detection

Subjects: [Remote Sensing](#)

Contributor: Yukun Wang , Mengmeng Wang , Zhonghu Hao , Qiang Wang , Qianwen Wang , Yuanxin Ye

Change detection (CD) stands out as a pivotal yet challenging task in the interpretation of remote sensing images. Significant developments have been witnessed, particularly with the rapid advancements in deep learning techniques. Nevertheless, challenges such as incomplete detection targets and unsmooth boundaries remain as most CD methods suffer from ineffective feature fusion.

change detection

remote sensing images

multi-scale progressive fusion

1. Introduction

The advance of satellite imaging technology has facilitated the acquisition of remote sensing images (RSIs). Change detection (CD) is the process of identifying changes in the ground within the same geographical area utilizing RSIs taken at two different times ^[1]. Due to its wide application in urban sprawl detection ^[2], urban green ecosystems ^[3], damage assessment ^[4], etc., CD as a fundamental and important task has increasingly gained attention in the remote sensing field.

During the early stages of CD research, numerous methods have been proposed by researchers ^{[5][6]}. For example, image difference was one of the earliest CD methods for subtracting bi-temporal images according to the corresponding pixels ^[7]. To address spurious changes and counter positional errors, a robust change vector analysis method was proposed by Thonfeld et al. ^[8], combining intensity information with the advantages of change vector analysis (CVA). Researchers have made substantial progress through extensive research on these traditional methods ^{[9][10][11]}. However, these traditional CD methods face new challenges with the increased spatial resolution of remote sensing images. On one hand, traditional CD methods are designed for medium- and low-resolution RSIs, resulting in poor performance when dealing with rich information in high-resolution RSIs ^[12]. On the other hand, these methods rely on handcrafted features that are sensitive to radiation differences and illumination changes ^{[13][14]}. Consequently, the application of traditional CD methods is limited in scope.

Recently, with the advent of the big data era, deep neural networks have demonstrated their strong feature extraction capabilities ^{[15][16]}, with the end-to-end advantages of convolutional neural networks (CNNs) being particularly notable. CNNs have been widely employed in CD tasks and have spawned a number of promising CD methods ^{[17][18]}. For example, Zhang et al. ^[19] integrated a CycleMLP block into a Siamese network, proposing an MLP-based method for CD. However, it is important to note that this method incurs a substantial inference time. Fang et al. ^[20] introduced a CD method that combines the UNet++ architecture with a Siamese network. This

method mitigates the loss of localization feature information by establishing a dense connection between the encoder and decoder.

Although the methods mentioned above have achieved performance results, they do not consider the characteristics of bi-temporal multi-scale features, thereby resulting in incomplete detection targets and limited accuracy of results. Inspired by the widely used multi-scale pyramid architecture for extracting multi-scale feature information in medical image segmentation [21], several methods have been proposed to address these problems by using multi-scale features [22][23][24]. For instance, Li et al. [23] proposed a multi-scale convolutional channel attention mechanism to generate detailed local features and integral global features. For capturing feature information on all scales, Xiang et al. [22] introduced a multi-receptive field position enhancement module incorporating convolutional layers with different kernel sizes. Despite the improvements achieved by the above methods through the incorporation of multi-scale features, they still exhibit certain shortcomings. On the one hand, these methods employ a simple concatenation strategy for fusing multi-scale features without considering the interaction between them. On the other hand, they extract multi-scale features after a simple feature fusion (i.e., feature difference) rather than employing bi-temporal multi-scale feature fusion. Consequently, the simple feature fusion often has restrictions that are not discriminative enough and result in unsmooth detection boundaries.

2. CNN-Based Methods

From the perspective of the fusion strategy, CNN-based methods can be further categorized into single-stream and two-stream methods [25]. In detail, single-stream methods take inspiration from semantic segmentation tasks. Researchers have proposed some approaches to image-level fusion strategies that match the semantic segmentation networks. For instance, Sun et al. [26] introduced conventional long short-term memory into Unet for CD. Peng et al. [27] employed bi-temporal images that had been concatenated into a UNet++ network. They further proposed a fusion strategy on multiple side outputs to improve the accuracy of results. Nevertheless, the independent feature characteristics of each bi-temporal image cannot be directly captured by single-stream CD methods based on semantic segmentation networks.

In contrast to single-stream, two-stream methods leverage the Siamese architecture, which consists of two streams that share weights to generate features of bi-temporal images. Most existing CD methods [20][28][29][30] adopt the Siamese architecture because it is appropriate for handling the input of RSIs. For instance, Dai et al. [28] introduced a building CD method that comprises a multi-scale joint supervision module and an improved consistency regularization module. Ye et al. [29] employed Siamese networks to propose a feature decomposition optimization reorganization network for CD. The edge and main body features were modeled using a feature decomposition strategy. Li et al. [31] proposed a lightweight CD method composed of three modules: a neighbor aggregation module (NAM), a progressive change identifying module (PCIM), and a supervised attention module (SAM), to improve the accuracy of results. Zhou et al. [32] introduced a context aggregation method utilizing a Siamese network. The multi-level features were fed into a context extraction module in this method, enabling the acquisition of long-range spatial-channel context features.

3. Transformer-Based Methods

Transformer-based methods, originally developed for natural language processing, are now being applied to encode bi-temporal images for CD. For example, Bandara et al. [33] introduced a CD method that combines a transformer with a Siamese architecture. This method introduced a transformer feature encoder to extract coarse and fine features with high and low resolution, respectively. Song et al. [34] introduced a progressive sampling transformer network (PSTNet) by using the excellent modeling ability of the transformer. In this method, the optimized tokens are iteratively mapped back to the original features to establish enhanced spatial connections in the spatial domain. Fang et al. [35] introduced a CD method, Changer, which uses a Siamese hierarchical transformer to extract multilayered features and then designs a flow-based dual-alignment fusion module to fuse the two branches' features. Zhang et al. [36] introduced a CD method that used a pure Swin transformer utilizing a Siamese network to extract long-term global features. However, transformer-based methods face limitations in terms of computational complexity and larger parameter sizes [37]. In addition, transformer-based methods often result in irregular boundaries in the results due to their disregard for the subtle details of shallow features.

4. Hybrid-Based Methods

Hybrid-based methods combine CNN and transformer architectures, which aim to improve feature extraction abilities [38]. For example, to couple the global and local features, Feng et al. [39] integrated a transformer and a CNN to design a CD method that was composed of an inter-scale feature fusion module and an intra-scale cross-interaction module, which were designed for obtaining discrimination feature maps and constructing spatial-temporal contextual information, respectively. To address the issues of blurred edges and neglect caused by sampling that is either too shallow or too deep, Song et al. [40] introduced a simple convolutional network and a progressive sampling CNN to generate fine and coarse features, respectively. Subsequently, a mixed-attention module was introduced to merge coarse and fine features. Finally, the results were generated by feeding the fused features into a transformer decoder. Chu et al. [41] proposed a dual-branch feature-guided aggregation network for CD. This method employs a dual-branch structure composed of a CNN and a transformer to extract both semantic and spatial features at various scales. However, in this method, the feature extractor is not only complicated but the network also has a large number of parameters. Tang et al. [42] introduced a W-shaped dual Siamese network (WNet) for CD. In this method, a deformable convolution was introduced into the CNN branch and transformer to mitigate the limited receptive fields and regular patch generation, respectively. Similarly, this method also possesses a significant number of parameters. Moreover, hybrid-based CD methods further require the design of a complicated fusion module to fuse the CNN features and token features, which are extracted from the CNN network and transformer network, respectively.

References

1. Lu, D.; Mausel, P.; Brondízio, E.; Moran, E. Change Detection Techniques. *Int. J. Remote Sens.* 2004, 25, 2365–2401.
2. Huang, X.; Zhang, L.; Zhu, T. Building Change Detection from Multitemporal High-Resolution Remotely Sensed Images Based on a Morphological Building Index. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 2014, 7, 105–115.
3. Shi, Q.; Liu, M.; Marinoni, A.; Liu, X. UGS-1m: Fine-Grained Urban Green Space Mapping of 31 Major Cities in China Based on the Deep Learning Framework. *Earth Syst. Sci. Data* 2023, 15, 555–577.
4. Gao, F.; Dong, J.; Li, B.; Xu, Q.; Xie, C. Change Detection from Synthetic Aperture Radar Images Based on Neighborhood-Based Ratio and Extreme Learning Machine. *J. Appl. Remote Sens.* 2016, 10, 046019.
5. Fang, H.; Du, P.; Wang, X.; Lin, C.; Tang, P. Unsupervised Change Detection Based on Weighted Change Vector Analysis and Improved Markov Random Field for High Spatial Resolution Imagery. *IEEE Geosci. Remote Sens. Lett.* 2022, 19, 6002005.
6. Wu, J.; Li, B.; Qin, Y.; Ni, W.; Zhang, H. An Object-Based Graph Model for Unsupervised Change Detection in High Resolution Remote Sensing Images. *Int. J. Remote Sens.* 2021, 42, 6209–6227.
7. Fung, T. An Assessment of TM Imagery for Land-Cover Change Detection. *IEEE Trans. Geosci. Remote Sens.* 1990, 28, 681–684.
8. 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.
9. Wu, C.; Du, B.; Zhang, L. Slow Feature Analysis for Change Detection in Multispectral Imagery. *IEEE Trans. Geosci. Remote Sens.* 2014, 52, 2858–2874.
10. Wang, M.; Han, Z.; Yang, P.; Zhu, B.; Hao, M.; Fan, J.; Ye, Y. Exploiting Neighbourhood Structural Features for Change Detection. *Remote Sens. Lett.* 2023, 14, 346–356.
11. 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.
12. Chen, H.; Wu, C.; Du, B.; Zhang, L.; Wang, L. Change Detection in Multisource VHR Images via Deep Siamese Convolutional Multiple-Layers Recurrent Neural Network. *IEEE Trans. Geosci. Remote Sens.* 2020, 58, 2848–2864.
13. Zhu, B.; Yang, C.; Dai, J.; Fan, J.; Qin, Y.; Ye, Y. R₂FD₂: Fast and Robust Matching of Multimodal Remote Sensing Images via Repeatable Feature Detector and Rotation-Invariant Feature Descriptor. *IEEE Trans. Geosci. Remote Sens.* 2023, 61, 5606115.

14. Ye, Y.; Zhu, B.; Tang, T.; Yang, C.; Xu, Q.; Zhang, G. A Robust Multimodal Remote Sensing Image Registration Method and System Using Steerable Filters with First- and Second-Order Gradients. *ISPRS J. Photogramm. Remote Sens.* 2022, 188, 331–350.
15. Ye, Y.; Tang, T.; Zhu, B.; Yang, C.; Li, B.; Hao, S. A Multiscale Framework with Unsupervised Learning for Remote Sensing Image Registration. *IEEE Trans. Geosci. Remote Sens.* 2022, 60, 5622215.
16. Tao, H.; Duan, Q.; An, J. An Adaptive Interference Removal Framework for Video Person Re-Identification. *IEEE Trans. Circuits Syst. Video Technol.* 2023, 33, 5148–5159.
17. Ye, Y.; Wang, M.; Zhou, L.; Lei, G.; Fan, J.; Qin, Y. Adjacent-Level Feature Cross-Fusion With 3-D CNN for Remote Sensing Image Change Detection. *IEEE Trans. Geosci. Remote Sens.* 2023, 61, 5618214.
18. Zhou, Y.; Feng, Y.; Huo, S.; Li, X. Joint Frequency-Spatial Domain Network for Remote Sensing Optical Image Change Detection. *IEEE Trans. Geosci. Remote Sens.* 2022, 60, 5627114.
19. Zhang, C.; Wang, L.; Cheng, S.; Li, Y. SUMLP: A Siamese U-Shaped MLP-Based Network for Change Detection. *Appl. Soft Comput.* 2022, 131, 109766.
20. Fang, S.; Li, K.; Shao, J.; Li, Z. SNUNet-CD: A Densely Connected Siamese Network for Change Detection of VHR Images. *IEEE Geosci. Remote Sens. Lett.* 2022, 19, 8007805.
21. Hu, Q.; Wang, D.; Yang, C. PPG-Based Blood Pressure Estimation Can Benefit from Scalable Multi-Scale Fusion Neural Networks and Multi-Task Learning. *Biomed. Signal Process. Control* 2022, 78, 103891.
22. Xiang, X.; Tian, D.; Lv, N.; Yan, Q. FCDNet: A Change Detection Network Based on Full-Scale Skip Connections and Coordinate Attention. *IEEE Geosci. Remote Sens. Lett.* 2022, 19, 6511605.
23. Li, J.; Hu, M.; Wu, C. Multiscale Change Detection Network Based on Channel Attention and Fully Convolutional BiLSTM for Medium-Resolution Remote Sensing Imagery. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 2023, 16, 9735–9748.
24. Jiang, K.; Zhang, W.; Liu, J.; Liu, F.; Xiao, L. Joint Variation Learning of Fusion and Difference Features for Change Detection in Remote Sensing Images. *IEEE Trans. Geosci. Remote Sens.* 2022, 60, 4709918.
25. Wen, D.; Huang, X.; Bovolo, F.; Li, J.; Ke, X.; Zhang, A.; Benediktsson, J.A. Change Detection From Very-High-Spatial-Resolution Optical Remote Sensing Images: Methods, Applications, and Future Directions. *IEEE Geosci. Remote Sens. Mag.* 2021, 9, 68–101.
26. Sun, S.; Mu, L.; Wang, L.; Liu, P. L-UNet: An LSTM Network for Remote Sensing Image Change Detection. *IEEE Geosci. Remote Sens. Lett.* 2022, 19, 8004505.

27. Peng, D.; Zhang, Y.; Guan, H. End-to-End Change Detection for High Resolution Satellite Images Using Improved UNet++. *Remote Sens.* 2019, 11, 1382.
28. Dai, Y.; Zhao, K.; Shen, L.; Liu, S.; Yan, X.; Li, Z. A Siamese Network Combining Multiscale Joint Supervision and Improved Consistency Regularization for Weakly Supervised Building Change Detection. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 2023, 16, 4963–4982.
29. Ye, Y.; Zhou, L.; Zhu, B.; Yang, C.; Sun, M.; Fan, J.; Fu, Z. Feature Decomposition-Optimization-Reorganization Network for Building Change Detection in Remote Sensing Images. *Remote Sens.* 2022, 14, 722.
30. Wang, M.; Zhu, B.; Zhang, J.; Fan, J.; Ye, Y. A Lightweight Change Detection Network Based on Feature Interleaved Fusion and Bistage Decoding. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 2024, 17, 2557–2569.
31. Li, Z.; Tang, C.; Liu, X.; Zhang, W.; Dou, J.; Wang, L.; Zomaya, A.Y. Lightweight Remote Sensing Change Detection with Progressive Feature Aggregation and Supervised Attention. *IEEE Trans. Geosci. Remote Sens.* 2023, 61, 5602812.
32. Zhou, F.; Xu, C.; Hang, R.; Zhang, R.; Liu, Q. Mining Joint Intraimage and Interimage Context for Remote Sensing Change Detection. *IEEE Trans. Geosci. Remote Sens.* 2023, 61, 4403712.
33. Bandara, W.G.C.; Patel, V.M. A Transformer-Based Siamese Network for Change Detection. In *Proceedings of the IGARSS 2022 IEEE International Geoscience and Remote Sensing Symposium, Kuala Lumpur, Malaysia, 17–22 July 2022*; pp. 207–210.
34. Song, X.; Hua, Z.; Li, J. PSTNet: Progressive Sampling Transformer Network for Remote Sensing Image Change Detection. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 2022, 15, 8442–8455.
35. Fang, S.; Li, K.; Li, Z. Changer: Feature Interaction Is What You Need for Change Detection. *IEEE Trans. Geosci. Remote Sens.* 2023, 61, 5610111.
36. Zhang, C.; Wang, L.; Cheng, S.; Li, Y. SwinSUNet: Pure Transformer Network for Remote Sensing Image Change Detection. *IEEE Trans. Geosci. Remote Sens.* 2022, 60, 5224713.
37. Song, L.; Xia, M.; Weng, L.; Lin, H.; Qian, M.; Chen, B. Axial Cross Attention Meets CNN: Bibranch Fusion Network for Change Detection. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 2023, 16, 32–43.
38. Song, X.; Hua, Z.; Li, J. LHDACT: Lightweight Hybrid Dual Attention CNN and Transformer Network for Remote Sensing Image Change Detection. *IEEE Geosci. Remote Sens. Lett.* 2023, 20, 7506005.
39. Feng, Y.; Xu, H.; Jiang, J.; Liu, H.; Zheng, J. ICIF-Net: Intra-Scale Cross-Interaction and Inter-Scale Feature Fusion Network for Bitemporal Remote Sensing Images Change Detection. *IEEE Trans. Geosci. Remote Sens.* 2022, 60, 4410213.

40. Song, X.; Hua, Z.; Li, J. Remote Sensing Image Change Detection Transformer Network Based on Dual-Feature Mixed Attention. *IEEE Trans. Geosci. Remote Sens.* 2022, 60, 5920416.
41. Chu, S.; Li, P.; Xia, M.; Lin, H.; Qian, M.; Zhang, Y. DBFGAN: Dual Branch Feature Guided Aggregation Network for Remote Sensing Image. *Int. J. Appl. Earth Obs. Geoinf.* 2023, 116, 103141.
42. Tang, X.; Zhang, T.; Ma, J.; Zhang, X.; Liu, F.; Jiao, L. WNet: W-Shaped Hierarchical Network for Remote-Sensing Image Change Detection. *IEEE Trans. Geosci. Remote Sens.* 2023, 61, 5615814.

Retrieved from <https://encyclopedia.pub/entry/history/show/123816>