

Application and Optimization of Deep Learning in Change Detection for High-Resolution Remote Sensing Imagery

Quan Qi, Yuhua Wang

Henan University of Technology, Henan, Zhengzhou, China

Abstract: With the continuous development of remote sensing technology and deep learning, change detection methods based on high-resolution remote sensing images are gradually evolving towards intelligence and high precision. Starting from the theoretical foundation of remote sensing image change detection, this paper systematically comprehends the technical framework and typical models of deep learning, and focuses on the analysis of its application modes in image alignment, feature extraction and bi-phasic analysis. In addition, the integration of multi-source remote sensing data and model adaptation are discussed with the idea of pixel-level and object-level modelling, which provides theoretical basis and methodological support for improving the accuracy and stability of change detection. The study shows that deep learning has a powerful characterisation capability and is an important development direction for change detection in remote sensing images in the future.

Keywords: Remote sensing technology; combined pixel level; remote sensing image.

1. Introduction

With the continuous development of remote sensing technology and the continuous improvement of satellite resolution, high-resolution remote sensing images have been widely used in the fields of land use monitoring, urban expansion analysis, and natural disaster assessment. The improvement of image resolution makes the details of features richer, but it also brings new challenges such as huge data volume and complex change features, and the traditional change detection methods have been difficult to meet the practical needs in terms of accuracy, robustness and automation level[1]. In recent years, deep learning, with its powerful feature extraction and representation capabilities, has achieved remarkable results in the field of computer vision and gradually expanded to the field of remote sensing image processing, especially in the change detection task, showing great potential. By constructing deep neural network models, complex spatial and semantic features in remote sensing images can be automatically learnt to effectively improve the accuracy and efficiency of change detection[2]. However, the direct application of deep learning methods to high-resolution remote sensing images still faces many problems, such as the model structure is not adapted to the characteristics of remote sensing data, insufficient training samples, high risk of overfitting, and limited model generalisation ability. Therefore, a systematic study on the application and optimisation of deep learning in remote sensing change detection not only has important theoretical value, but also has outstanding engineering application significance[3].

2. The Basis of Change Detection of High-Resolution Remote Sensing Images

2.1. Definition and characteristics of high-resolution remote sensing images

High-resolution remote sensing images usually refer to

remote sensing data with a spatial resolution better than 1 metre, which can provide rich feature details and boundary information, and are widely used in fields such as urban planning, resource investigation and disaster assessment. Its distinctive features include: (1) clear image details, which can distinguish small-scale features; (2) large data volume and high computational complexity; (3) rich information dimensions, including spectral, spatial and textural features[4]; and (4) sensitive to the influence of noise and environmental factors. High-resolution remote sensing data provide the basis for fine feature identification and change detection, but also put forward higher requirements for image processing technology, especially in feature extraction, alignment accuracy and change difference identification, which need to rely on more advanced algorithms to solve[5].

2.2. Basic principle and classification of change detection

Change detection refers to the process of identifying the change information of surface targets through comparative analysis between remote sensing images of the same region obtained in different periods. Its basic principles include image alignment, radiation normalisation, difference analysis and result interpretation. According to the different processing methods, the change detection methods can be divided into: based on image difference method (e.g., image difference, ratio, principal component analysis), based on classification method (e.g., post-classification comparison, decision-level fusion) and based on deep learning method. According to the spatial scale of the research object and application requirements, the change detection task can be further divided into pixel-oriented, object-oriented and semantic-oriented detection methods[6].

2.3. Main Challenges of Change Detection in Remote Sensing Images

Although high-resolution imagery brings rich feature information, it faces the following challenges in change detection applications: (1) image alignment errors have a

significant impact on detection accuracy; (2) seasonal variations, lighting differences and other non-featured factors are prone to pseudo-changes; (3) blurred boundaries of complex features increase the difficulty of change discrimination; (4) data redundancy and computational resource consumption of high-resolution imagery are high; and (5) Model training relies on a large amount of labelled data, and remote sensing data annotation is costly[7].

2.4. Common change detection evaluation index system

The effectiveness of change detection results needs to be verified by a scientific evaluation index system. Commonly used indicators include overall accuracy (OA), Kappa coefficient, detection accuracy (PA), user accuracy (UA), false positives rate (FPR) and missed positives rate (FNR). Among them, the overall accuracy measures the proportion of correct classifications, the Kappa coefficient reflects the difference between the results and the random consistency; PA and UA are used to assess the ability of a certain type of change to be correctly recognised and the likelihood of misrecognition, respectively.

3. Theoretical Foundations of Deep Learning Techniques

3.1. Development and basic principles of deep learning

Deep Learning (Deep Learning) originated from the study of artificial neural networks, is a machine learning method based on multi-layer nonlinear transformation, and has made breakthroughs in recent years in the fields of image recognition, speech recognition, natural language processing and remote sensing change detection. Its core idea is to automatically extract multilevel, high-abstraction feature representations by building a deep network structure, so as to effectively deal with the bottleneck of traditional machine learning in feature engineering.

3.2. Convolutional Neural Network (CNN) Structure Principle

Convolutional Neural Network (CNN) is one of the most widely used deep learning models in the field of image processing, which is especially suitable for feature extraction and classification of remote sensing images. The structure of CNN is mainly composed of convolutional layer, pooling layer, activation function layer and fully connected layer, in which the convolutional layer achieves the high efficiency of local spatial structure through the mechanism of local sensing field and weight sharing. Among them, the convolutional layer achieves efficient modelling of local spatial structure through local receptive field and weight sharing mechanism, which greatly reduces the number of parameters and improves the learning efficiency.

3.3. Introduction of typical deep learning models

Several deep learning models are widely adopted in remote sensing image change detection, especially Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), U-Net, ResNet, DenseNet, and Transformer-based architectures. CNN, as a basic model, is mainly used for image classification and local feature extraction, while U-Net, as a

typical U-Net, as a typical semantic segmentation network structure, can achieve pixel-level change region identification, and its structure adopts a symmetric encoder-decoder architecture, which fuses the low-level details and high-level semantic information through jump connections, and is widely used in the generation of binary change maps. ResNet introduces the residual connection mechanism, which effectively mitigates the gradient disappearance and degradation in the deeper network, and enables the model to learn the more complex feature expressions more deeply, and improves the detection of changes in remote sensing imagery.

3.4. Training Mechanism and Optimisation Strategy of Deep Learning Models

Deep learning model training is the process of optimising the objective function to minimise the prediction error, and the back propagation algorithm combined with the gradient descent method is usually used to update the parameters. In the training process, the design of the loss function is crucial, and the commonly used ones are Cross Entropy Loss (Cross Entropy), Dice Loss, IoU Loss, etc., which are suitable for tasks such as classification, segmentation and target detection, respectively. In order to improve the training effect and convergence speed, the choice of optimiser is also very critical, such as SGD, Adam, RMSProp, etc. have their own advantages and disadvantages, among which Adam is widely used in remote sensing image modelling due to its adaptive learning rate mechanism. Meanwhile, data enhancement techniques (e.g., rotation, cropping, colour perturbation) are widely used to expand the training set and improve the generalisation ability of the model. Regularisation means such as Dropout, Batch Normalization, and L2 regular terms are used to prevent overfitting and enhance the robustness of the model. In addition, learning rate scheduling strategies (e.g., cosine annealing, learning rate decay) can effectively control the model convergence speed and performance fluctuations during the training process. In remote sensing change detection applications, due to the scarcity of labelled data, it is also necessary to combine migration learning and semi-supervised learning methods to improve the learning ability of the model under small sample conditions.

4. Patterns of Deep Learning Application in Remote Sensing Change Detection

4.1. Image Alignment and Difference Feature Extraction Methods

In remote sensing image change detection, image alignment and difference feature extraction are the prerequisite and foundation for multi-temporal image analysis. Image alignment refers to the process of aligning remote sensing images acquired by different time phases, different viewpoints or different sensors to the same spatial reference system, and commonly used methods include geometric alignment based on feature point matching, dense alignment based on grey scale correlation, and multiscale alignment techniques that integrate geographic information. With the development of deep learning, end-to-end image alignment models (e.g. based on CNN or Siamese network) can automatically extract spatial features and complete high-precision alignment, providing favourable conditions for subsequent change region extraction. In terms of difference

feature extraction, traditional methods rely on image differencing, change vector analysis (CVA), etc., but have limited ability to identify complex feature changes.

4.2. Deep Learning Architecture for Dual-Temporal Image Analysis

The task of detecting changes in dual-time-phase remote sensing images is essentially an automated identification of the differences between images of different periods in the same region, so it is necessary to build a deep learning architecture that can handle image pairs. Typical model architectures include Siamese network based on shared weights, Difference Network, fusion encoder-decoder structure (e.g., improved U-Net), and temporal graph neural network (ST-GCN). Among them, Siamese network inputs two temporal images through two branches at the same time, extracts high-level semantic features and determines the change region through feature difference calculation, which can make full use of temporal consistency and improve the model's ability to perceive subtle changes. The fused architecture combines the two input images and constructs the change map through the multi-scale feature extraction module and the up-sampling decoding module, thus achieving end-to-end training and pixel-level segmentation results.

4.3. Pixel- and object-oriented detection model analysis

In remote sensing image change detection, pixel-oriented and object-oriented detection models are two major analysis methods, each with its own advantages and applicable scenarios. Pixel-oriented methods generate change probability maps or binary change maps by estimating the change probability of each pixel in the image, which is characterised by high computational accuracy and detailed spatial resolution. Deep learning models such as FCN and U-Net have superior performance in pixel-level segmentation, and are able to obtain rich spatial and semantic information through mechanisms such as multi-scale context fusion and jump connections, and achieve refined change boundary inscription. However, pixel-oriented methods are more sensitive to noise and local inconsistencies, which may lead to 'pretzel noise' errors. In contrast, object-oriented methods form a number of object units with semantic information through image segmentation or clustering, and then identify and classify changes at the object level, which is suitable for high-resolution remote sensing images and complex surface structures.

4.4. Research on fusion of multi-source remote sensing data and deep model adaptation

With the development of remote sensing technology, the types of data acquired show a trend of multi-source, including optical images, synthetic aperture radar (SAR), LiDAR, and multi-spectral and hyperspectral images. These data have significant differences in spatial resolution, spectral characteristics, observation methods, etc. How to fuse the multi-source information and adapt the deep learning model has become a key issue in change detection. Multi-source data fusion mainly includes data-level fusion, feature-level fusion and decision-level fusion, among which feature-level fusion

is the most widely used in deep learning frameworks, and cross-modal information integration is achieved by inputting different source data into multiple sub-networks to extract features and then fusing them (e.g., channel splicing, attentional weighting, feature alignment, etc.). In the face of significantly different data sources, such as SAR and optical images, the researchers constructed a multimodal neural network, graph convolutional network (GCN), and Transformer fusion architecture to improve the model adaptability and discriminative ability. In addition, in order to enhance the robustness of the model to different source data, techniques such as data normalisation, domain adaptive training and migration learning are often adopted to maintain stable performance of the model when dealing with data heterogeneity.

5. Conclusion

This paper focuses on the change detection problem of high-resolution remote sensing images, and discusses the theoretical mechanism and practical path of the deep learning method in depth. It is found that the deep neural network can effectively mine the spatio-temporal features in the image to improve the accuracy and efficiency of change identification. Multi-source data fusion and structure optimisation further enhance the robustness and generalisation ability of the model. Although still facing challenges such as strong sample dependence and insufficient interpretability, deep learning will play an increasingly important role in the field of remote sensing change detection through model improvement and multimodal information integration, providing technical support for applications such as land use monitoring, environmental assessment and disaster emergency response.

References

- [1] WANG Q, ZHANG X, CHEN G, et al. Change detection based on Faster R-CNN for high-resolution remote sensing images [J]. *Remote Sensing Letters*, 2018.
- [2] WANG M, ZHANG H, SUN W, et al. A Coarse-to-Fine Deep Learning Based Land Use Change Detection Method for High-Resolution Remote Sensing Images [J]. *Remote Sensing*, 2020.
- [3] XING, HUAQIAO, ZHU, et al. A comparative study of threshold selection methods for change detection from very high-resolution remote sensing images [J]. *Earth Science Informatics*, 2022.
- [4] HUANG L, AN R, ZHAO S, et al. A Deep Learning-Based Robust Change Detection Approach for Very High Resolution Remotely Sensed Images with Multiple Features [J]. *Remote Sensing*, 2020.
- [5] LARABI M E A, CHAIB S, BAKHTI K, et al. High-resolution optical remote sensing imagery change detection through deep transfer learning [J]. *Journal of Applied Remote Sensing*, 2019.
- [6] WANG H, QI J, LEI Y, et al. A Refined Method of High-Resolution Remote Sensing Change Detection Based on Machine Learning for Newly Constructed Building Areas [J]. *Remote Sensing*, 2021.
- [7] JIANG H, PENG M, ZHONG Y, et al. A Survey on Deep Learning-Based Change Detection from High-Resolution Remote Sensing Images [J]. *Remote Sensing*, 2022.