Research on change detection of high resolution remote sensing image based on U-type neural network

Congyi Li$^{1,2}$, Xiangbing Kong$^{3,4}$, Fuling Wang$^3$, Yinan Wang$^2$ and Mengxuan Zhang$^{1,2}$

$^1$School of Surveying and Land Information Engineering, Henan Polytechnic University, Jiaozuo, Henan 454003, China;
$^2$Yellow River Institute of Hydraulic Research, Zhengzhou Henan 450003, China;
$^3$Henan Soil and Water Conservation Monitoring Central Station, Zhengzhou Henan 45000, China

$^4$Email: kongxb_whu@foxmail.com

Abstract. The change detection of high-resolution remote sensing images has always been a research difficulty and hot spot, which has a great application demand in land use change and ecological environment monitoring. U-type neural network can train a better model with less samples, and has good experimental effect in medical image segmentation, but it is rarely used in remote sensing image change detection research. In this paper, the U-type neural network is used to detect the change spots in GF-1 image of Yuzhou City, Henan Province, and compared with FCN, SegNet and Siamese-CNN. Experimental results show that the F1 score of U-type neural network model are 0.699 and 0.66 respectively, which are better than the other three methods, and the omission is lower, which is closer to the label map. Therefore, it is feasible to use U-type neural network for change detection of high-resolution remote sensing images, and it can have high detection accuracy.

1. Introduction

With the successful launch of a series of high spatial resolution remote sensing satellites such as QuickBird, IKONOS, ZY-3, GF-1, etc., the characteristic distribution and spatial correlation between ground objects are clearly presented. The detection of land cover change through high-resolution remote sensing images has become an efficient and accurate technical means [1]. The change detection of remote sensing image is to determine whether the ground features in the region have changed through two (or more) images acquired in different periods of the same geographical area, combined with relevant geographic data and remote sensing imaging mechanism[2]. For remote sensing image change detection technology, scholars at home and abroad have proposed a large number of algorithms, such as algebraic method [3], time series analysis method [4], object-oriented method [5, 6]. Although these methods have been widely used in the fields of land cover / land use monitoring [7], urban expansion [8], ecosystem monitoring [9-11], disaster monitoring [12], with the aggravation of land change speed, the deepening of environmental complexity and the increasing diversity of remote sensing data, the traditional methods can no longer meet the needs of change detection [13].

In recent years, deep learning has made great progress in image classification, semantic segmentation and object detection with its powerful feature extraction ability. In 2015, long proposed the full convolution neural network (FCN) [14], compared with the traditional convolutional neural
network (CNN). After multi-layer convolution and pooling operation, FCN does not use full connection layer to construct feature vector to predict image, but uses deconvolution operation to upsample the feature map, so that the output image and the input image have the same resolution, so as to achieve the pixel level semantic segmentation of the image. Although the multi feature fusion of short links is used in FCN network, the spatial consistency between pixels and the global context information of the image are not fully considered in the prediction of each pixel, which leads to the fuzzy edge of the final prediction result and insufficient segmentation accuracy. The subsequent U-type neural network (UNet) is improved on the basis of FCN. Compared with the short connection used in FCN, UNet uses jump connection to fuse low-dimensional and high-dimensional features at the decoding end, so as to improve the problem of blurred edge details [15], and has achieved good experimental results in the field of medical image segmentation, but it is rarely used in high-resolution remote sensing images Research on image change detection. Therefore, this paper attempts to use the UNet model to realize the change detection of remote sensing images.

2. U-type neural network structure

UNet [15] is a segmented network proposed by Ronneberge and others in the 2015 ISBI challenge competition. It is a symmetric U-type structure, and the network structure is shown in Figure 1.

![Figure 1. Architecture of UNet.](image)

The contraction path on the left side of the network is composed of four convolution layers. Each convolution layer is composed of two convolution cores with the size of $3 \times 3$. The adjacent convolution layers are downsampled by the maximum pooling operation with the size of $2 \times 2$ each time. The main function of the contraction path is to extract the features of the input image through multiple convolution and pooling operations, and finally form a high-dimensional feature map. The expansion path at the right end is a decoding structure composed of four upper sampling layers. The up sampling process is carried out by deconvolution, which can avoid the problem of feature loss in the transmission process. Jump join fuses the deep and shallow feature maps in the band dimension, and retains more dimensional location information, thus improving the problem of blurred image edge details.

3. Experiment and analysis

3.1. Experimental data

The experimental area is Yuzhou City, Henan Province, and the data used are GF-1 remote sensing images in 2016 and 2017. Through geometric correction, radiation correction and other pretreatment, the selected experimental sample are two training samples are 10000 $\times$ 10000 pixels in size, which are selected from the same spatial position of the two images; the validation samples are 960 $\times$ 960 pixels in size, which are also the same spatial position of the two images. Combined with field investigation
and visual interpretation, the black-and-white binary label map is shown in Figure 2, in which black represents unchanged area and white represents changed area.

![Training samples of Yuzhou City in 2016](image1)
![Training samples of Yuzhou City in 2017](image2)
![Label](image3)

**Figure 2.** Training samples and label of two temporal images.

### 3.2. Experimental environment and model parameter setting

The experimental environment is i7-8700 processor, NVIDIA GEFORCE GTX1080 graphics card, 16G memory, GPU acceleration library using CUDA-9.2. The framework of deep learning is PyTorch (https://pytorch.org/).

Before training the UNet model, we need to set the parameters. The parameter optimizer is ADMA function, learning rate is used to adjust the learning progress of the model, the initial learning rate is $1 \times 10^{-4}$; the learning rate change index (gamma) is used to adjust the change rate of learning rate. In this experiment, gamma is set to 0.1; when the loss value of verification set is no longer reduced, the learning rate is attenuated, and the reduced learning rate is equal to the product of the initial learning rate and the learning rate change index. The loss function of model training adopts the category cross entropy loss function. Due to the imbalance between the number of changed samples and the number of unchanged samples, Maxwell et al [16] showed that the semantic segmentation model based on deep convolutional neural network is sensitive to the balance of positive and negative samples, so it needs to be weighted. By counting the number of changing pixels and unchanging pixels, its weight was set to 1:9 to balance the positive and negative samples.

### 3.3. Model training and testing

In order to avoid the problem of computer memory overflow in the training of U-type neural network, the remote sensing images of training samples are cut into $960 \times 960$ small image blocks for training. In order to enhance the generalization ability of the model, the training samples were rotated by 90°, 180° and 270° by data enhancement method, and the training samples were expanded by vertical and horizontal rotation methods.

At the input end of the U-type neural network, the remote sensing images of two phases and three channels are superimposed to form a six channel remote sensing image for input. The input image is convoluted and pooled for many times through the contraction path on the left side of the network to extract the features of the input image. In order to accurately predict whether each pixel changes or not, the feature map needs to be restored to the size of the input image. Therefore, in the right expansion path of the network, the feature map is upsampled by deconvolution operation, and the softmax classifier is used to predict the feature map. Finally, a change detection image including change and unchanged classification is obtained.

### 3.4. Experimental results and analysis

In order to verify the effectiveness of the U-type neural network in the change detection of high-resolution remote sensing images, FNC [14] SegNet [17] and Siamese-CNN [18] are used to carry out change detection experiments on two phases of images, and the effectiveness is compared and analyzed. Through the statistical analysis compared with the label chart, the quantitative evaluation indexes [19,20] such as precision, recall, omission and F1 score of the experimental results of the four change detection methods were obtained.
The higher Precision, Recall and F1 score of the evaluation index, the better the detection effect; the lower the omissions, the better the detection effect. From the perspective of quantitative analysis of area 1, as shown in Table 1, compared with SegNet, FCN and Siamese-CNN, UNet method is superior to the other three methods in precision, recall and F1 score, in which the precision rate, recall rate and F1 score are 5.4%, 4.1% and 1.9%, 3.6%, 13.4% and 10.9%, 4.5%, 9% and higher respectively. It can be seen from Figure 3 that the change patterns detected by the three methods are taken as an example. The experimental results of SegNet method have fuzzy edge details, FCN method has obvious stitching traces, and Siamese-CNN has many missing change patches. Relatively speaking, the change patterns detected by UNet method are close to the label map (see the yellow box area for details); however, the Red Square is not detected by the three methods. The road in the box is detected. In general, the contour information of the change pattern detected by the UNet method is more complete, and the missing and wrong detection spots are less, which is closer to the label map. It can be seen from Table 2 that the precision rate, recall rate and F1 score of UNet method are higher than those of the other three methods compared with SegNet, FCN and Siamese-CNN. The F1 score is 10.6%, 4.3% and 9.4% higher respectively, and the missed detection rate is lower than the other three methods. It can be seen from Figure 4 that there are false detections in all four methods. In contrast, SegNet, FCN and Siamese CNN all have large areas of error detection, while the UNet method has a smaller error detection area, and the change detection map obtained is closer to the label map (see the yellow box for details).
4. Conclusions
In this paper, the semantic segmentation model based on deep learning is used to detect the change of high-resolution remote sensing images, and the automatic extraction of the change regions of the two images is realized. The experimental results show that, compared with the three traditional change detection methods, the semantic segmentation model based on deep learning is superior to the traditional change detection methods in Precision, Recall, Omission and F1 score. Compared with the other two models, UNet model also has better performance in quantitative evaluation indicators such as Omission and F1 score, and is closer to the label. However, the three semantic segmentation models still have the phenomenon of omission and error detection. Therefore, in the research of high-resolution remote sensing image change detection, how to use deep learning method to extract effective features, reduce false changes and improve the accuracy of change detection is the main direction of future research.

Acknowledgements
This paper was supported by the National Science Foundation of China (61501200), the National Key R&D Program of China (2017YFC0504501), the Water Conservancy Science and Technology Project of Henan Province (GG201942, GG201829), and the and the R&D Project of YRIHR(HKY-YFXM-2020-02).

References
[1] Zhang L P, WU C 2017 Advance and Future Development of Detection for Multi-temporal Remote Sensing Imagery [J] Acta Geodaetica et Cartographica Sinica 46(10) 1447-1459
[2] Bruzzone, Lorenzo Bovolo, Francesca 2013 A Novel Framework for the Design of Change-Detection Systems for Very-High-Resolution Remote Sensing Images [J] Proceedings of the IEEE 101(3) 609-630
[3] Terry L. Sohi 1999 Change analysis in the United Arab Emirates: An investigation of techniques [J] Photogrammetric Engineering & Remote Sensing 65(4) 475-484
[4] Zhao Z M, Meng Y, Yue Z A, et al 2016 Research progress of remote sensing time series image change detection [J] Journal of remote sensing 020(005) 1110-1125
[5] Li L, Shu N, Wang Y 2011 Remote sensing image change detection based on image patch using normalized mutual information [J] Remote sensing information 000(6) 18-22
[6] Baudouin Desclée, Patrick Bogaert, Pierre Defourny 2006 Forest change detection by statistical object-based method [J] Remote Sensing of Environment 102(12) 1-11
[7] Jin S M, Yang L M, Zhu Z, et al 2017 A land cover change detection and classification protocol
for updating Alaska NLCD 2001 to 2011. [J] Remote Sensing of Environment 195 44-55
[8] George Xian, Collin Homer 2010 Updating the 2001 National Land Cover Database Impervious Surface Products to 2006 using Landsat Imagery Change Detection Methods. [J] Remote Sensing of Environment 114(8) 1676-1686
[9] Ruth Sonnenschein, Tobias Kuemmerle, Thomas Udelhoven, et al 2011 Differences in Landsat-based trend analyses in drylands due to the choice of vegetation estimate. [J] Remote Sensing of Environment 115(6) 1408-1420
[10] Andrew J, Elmore, John F, et al 2000 Quantifying Vegetation Change in Semiarid Environments: Precision and Accuracy of Spectral Mixture Analysis and the Normalized Difference Vegetation Index. [J] Remote Sensing of Environment 73(1) 87-102
[11] Song C Q, Huang B, Ke L H, et al 2014 Remote sensing of alpine lake water environment changes on the Tibetan Plateau and surroundings: A review. [J] ISPRS Journal of Photogrammetry and Remote Sensing 92 26-37
[12] Brunner D, Lemoine G, Bruzzone L 2010 Earthquake Damage Assessment of Buildings Using VHR Optical and SAR Imagery. [J] IEEE Transactions on Geoscience & Remote Sensing 48(5) 2403-2420
[13] Daudt R C, Le Saux B, Boulch A 2018 Fully Convolutional Siamese Networks for Change Detection. [C]// 2018 25th IEEE International Conference on Image Processing (ICIP) 4063-4067
[14] Long J, Shellhamer E, Darrel T 2015 Fully Convolutional Networks for Semantic Segmentation. [C]// Proceedings of the IEEE conference on computer vision and pattern recognition 3431-3440
[15] Ronneberger O, Fischer P, Brox T 2015 U-Net: Convolutional Networks for Biomedical Image Segmentation. [C]// International Conference on Medical image computing and computer-assisted intervention 9351 234-241
[16] Aaron E. Maxwell, Timothy A. Waener, Fang F 2018 Implementation of machine-learning classification in remote sensing: an applied review. [J] International Journal of Remote Sensing 39(9) 2784-2817
[17] Badrinarayanan V, Kendall A, Cipolla R 2017 Segnet: A deep convolutional encoder architecture for image segmentation. [J] IEEE transactions on pattern analysis and machine intelligence 39(12) 2481-2495
[18] Chopra S, Hadsell R, LeCun Y 2005 Learning a similarity metric discriminatively, with application to face verification. [C]// Proceedings of 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition 539-546
[19] T. Lei, D. Xue, Lv Z, et al 2018 Unsupervised Change Detection Using Fast Fuzzy Clustering for Landslide Mapping from Very High-Resolution Images. [J] Remote Sensing 10(9) 1381
[20] T. Lei, Y. Zhang, Lv Z, et al 2019 Landslide Inventory Mapping From Bitemporal Images Using Deep Convolutional Neural Networks. [J] IEEE Geoscience and Remote Sensing Letters 1-5