Building Change Detection on High-Resolution Imagery with a Multi-task Semantic Change Detection Method

Zhipan Wang[1](✉), Xiang Liao[2], Huan Liu[3], and Hua Zhang[3]

[1] Sun Yat-Sen University, Shenzhen 528406, Guandong Province, China
wang749195@outlook.com
[2] MagicCubesat Technology Co.LTD., Shenzhen, Guandong Province, China
[3] Changsha University of Science and Technology, Changsha, Hunan Province, China

Abstract. Building change detection from high-resolution images plays an important role in landcover classification and the analysis of urban spatial information. With the development of deep learning (DL) in remote sensing, DL has become the mainstream method in the building change detection field. However, the existing building change detection methods still only can detect building change, but there are few studies investigating whether the building is increased or decreased, in other words, they can only capture the spatial change of buildings, but can't get the temporal change of buildings. To solve this problem, this manuscript proposes a new simple but effective multi-task learning building change detection method, which can capture building increase or decrease on high-resolution images at the same time. First, we modify the DenseNet121 as our backbone to extract deep semantic features, mainly because of DenseNet has a strong feature extraction ability, but the computation cost is relatively low. Second, we design a differencing structure to detect building increases or decreases at the same time with a new loss function, namely Tanimoto loss. A large image with very high resolution is selected to validate the effectiveness and robustness of our method, on the one hand, the result indicates that our proposed building change detection method can detect building increases or decreases accurately, and the quantitative precision shows that the Accuracy is over 90%, on the other hand, the model design idea of our building change detection model, which can provide some new insights to other research fields, such as object detection on high-resolution images, change detection of other objects.

Keywords: Building Extraction · Multi-task Learning · High-resolution Images

1 Introduction

Using remote sensing images to monitor large-area building change is an effective way for developing countries, such as illegal building construction monitoring. Deep learning methods, get an impressive result than traditional machine learning methods due to their strong self-learning ability [1]. Thus, building change detection with deep learning methods has become the mainstream method [2]. However, the full application ability of deep learning was not fully investigated.

H. P. Urbach and H. Jiang (Eds.): ISSOIA 2022, SPPHY 295, pp. 38–46, 2023.
https://doi.org/10.1007/978-981-99-4098-1_5

Building change detection can be divided into two categories, binary change detection (BCD) and semantic change detection (SCD) [3]. Binary change detection, which can only detect building change in bitemporal images, the whole model structure is similar to the semantic segmentation model, that is considers the change detection task as a pixel-level or object-level segmentation task [4]. SCD method, can not only detect building change but also obtain segmentation results. A classical SCD method is the semantic feature-constrained change detection method (SFCCD) [3], it can get three product types simultaneously, they are the former image segmentation result, the latter image segmentation result, and the change detection result.

The training dataset can influence the final detection accuracy heavily. There are excellent open-sourced datasets for building change detection, such as NJDS [3] and WBDS [4], Many of them are labeled from Google Earth images or other commercial satellite images. However, there are some limitations in these datasets. For instance, building increase or decrease information cannot be obtained. Building increases and decreases are important for illegal building construction investigations. Consequently, we need to develop a method, which can get the building to increase and decrease. In this paper, we develop a building semantic change detection method, which was trained based on the existing building semantic change detection dataset, such as NJDS [3] and WBDS [4]. Moreover, in order to help other researchers to investigate building change, we have developed an easy-used building change detection software, namely, Easy-Building Change.

The main contribution of this paper is as follows:

1. A multi-task building semantic change detection model is proposed, which can obtain building increase and decrease simultaneously.
2. The principle of our multi-task semantic change detection, can provide new insight ideal for other tasks, such as forest change detection, and road change detection.

2 Materials and Methods

2.1 Data

In this manuscript, we use WBDS dataset to train our building semantic change model. First, we clip the large image into small blocks, such as a small image with fixed size (512 $\times$ 512 pixels), then we mix these small patches, and split the whole dataset into three sub-datasets, they are train dataset, valid dataset, and test dataset, their ratio is 4:4:3, finally, we use the three sub-datasets to train our deep learning model. The detailed information about this dataset is as follows (Table 1):

Table 1. Building change detection dataset comparison

Open dataset	Spatial resolution	Number of samples (512 × 512 pixels)	Data description
NJDS [3]	0.3	617	In their paper, this dataset is open-sourced, but we cannot download their dataset
WBDS [4]	0.2	1902	This dataset was generated by Wuhan University, it is open-sourced and free to use

A large-area study area is chosen to validate the effectiveness of our deep learning method. The two different temporal images are processed from the SuperView satellite, which is captured in Jiangsu Province, China, and the spatial resolution of the two images is 0.5 m, and the image size is 16978 × 14955 pixels. The RGB true color images are as follows (Fig. 1):

(a) (b)

Fig. 1. Study area. (a) 2014 image. (b) 2022 image.

2.2 Change Detection Based on Multi-task Learning

Detecting building increase and decrease at the same time, which can be defined as a multi-class classification task [5]. Though it is simple and easy to accomplish, the detection accuracy of the relatively small number category may be unsatisfactory. Consequently, in this paper, we define building increase and decrease detection as a multi-task learning question, the whole architecture is as follows:

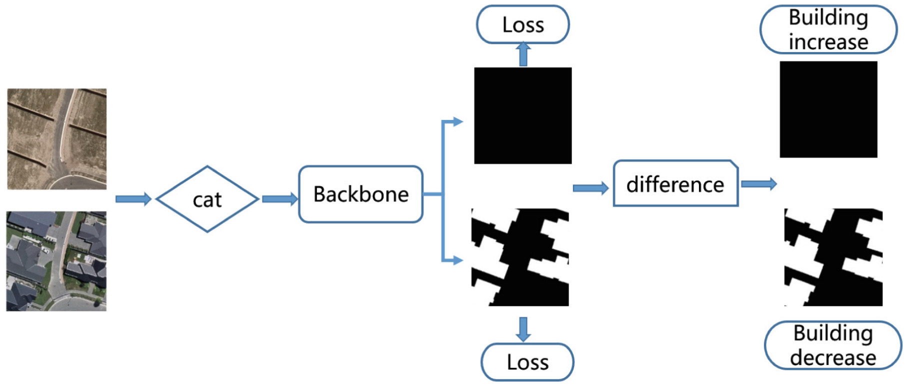

Fig. 2. Building change detection model

There are two main parts in our deep learning model, the backbone, and the semantic change detection head. DenseNet121 [6] is chosen as our backbone to extract deep semantic features, though there are many other excellent deep feature models, such as ResNet101 [7], ResNext101 [8], DenseNet101 and so on, the computation efficiency is not better. Considering the balance between efficiency and accuracy, we use DenseNet121 as the backbone. Then, we use a difference operation to get building increase or decrease, this operation is very simple, it is equal to basic math "subtract" operation.

The most important part of our building change detection model is the semantic change detection head, inspired by YOLOX [9], we use a decoupled head to detect building semantic change. When the abundant deep semantic feature of the building is extracted, two independent convolution module is utilized to extract building increase and decrease, respectively. From Fig. 2, the whole architecture is a multi-task learning model, it is simple and effective.

2.3 Loss Function

The loss function plays an important role in the deep learning method. Cross-entropy [10] is the most popular loss function in semantic segmentation [11], object detection, or instance segmentation. However, the classical cross-entropy function cannot solve the class-unbalance problem, in order to address this question, weighted cross-entropy is proposed [12], which is an effective and robust way to alleviate unbalanced class. Other excellent loss functions, such as focal loss [13], and OHEM [14], also give an interesting but effective way to help solve the class-unbalanced problem. Nonetheless, most of them need hyper-parameter setting, in other words, more prior knowledge is needed. Therefore, we introduce a new loss function, namely, Tanimoto loss [15], which can optimize classification and prediction tasks at the same time, and it also doesn't need any other parameter setting, moreover, small objection detection or segmentation with Tanimoto loss shows a robust result. The detailed information can be seen in [15].

2.4 Accuracy Assessment

We quantitatively evaluate the model performance using the three metrics, Accuracy, Recall, and F1-score [3], whose formulas are:

$$\text{Accuracy: } accuracy = (TP + TN)\,(Tp + TN + FP + FN) \tag{1}$$

$$\text{Recall: } recall = TP/(TP + FN) \tag{2}$$

$$\text{F1 - score: } F1 = 2TP/(2TP + FN + FP) \tag{3}$$

Here, TP indicates that the predicted result is a change pixel and the ground truth pixel is also a change pixel. TN indicates that the predicted pixel is the background value and the ground truth pixel is also the background value. FP indicates that the pixel in the ground truth image is changed but is predicted as the background value. FN indicates that the pixel in the ground truth image is the background value but is predicted as changed.

2.5 Results

The building change detection result of the study area can be seen in the following figure (Fig. 3).

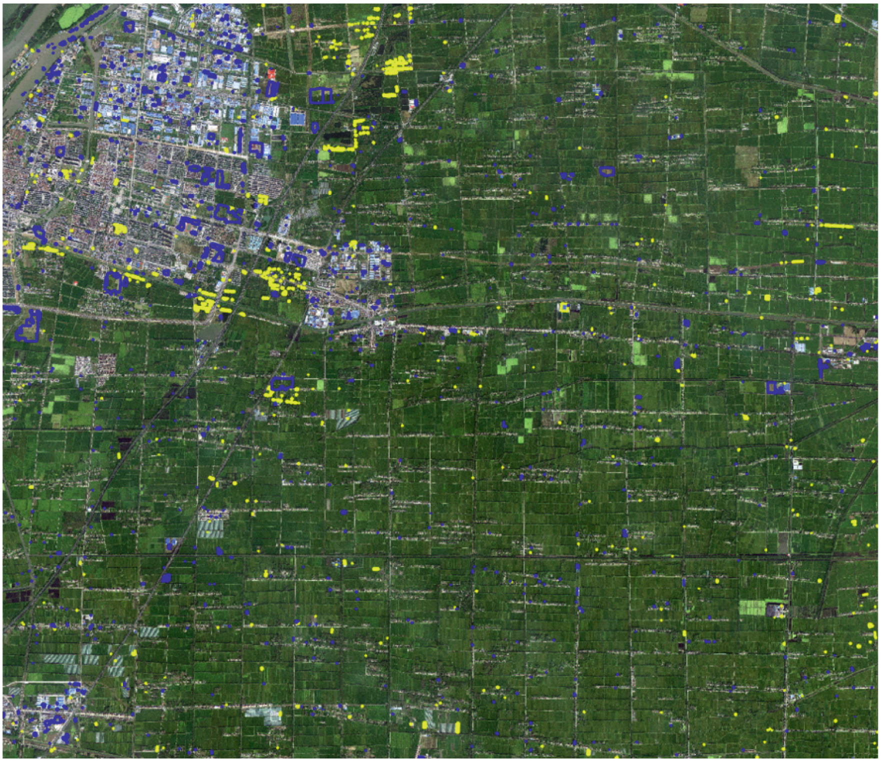

Fig. 3. Building change detection result. (The blue color and yellow color are building increase, building decrease, respectively)

The above figure indicates that building increase usually appears in the center of the city, while building decrease appears in the edge of the city or suburbs. In addition to the upper-left of the whole study area, building increases or decreases show a uniform distribution in other regions. A few small patch areas are chosen to describe the detailed information about building semantic change (Fig. 4).

We can see that, our change detection model can detect building increase and decrease accurately, whether it is a small building change or a large building change, our method can detect accurately outline the changed boundary.

The quantitative accuracy assessment is in Table (Table 2).

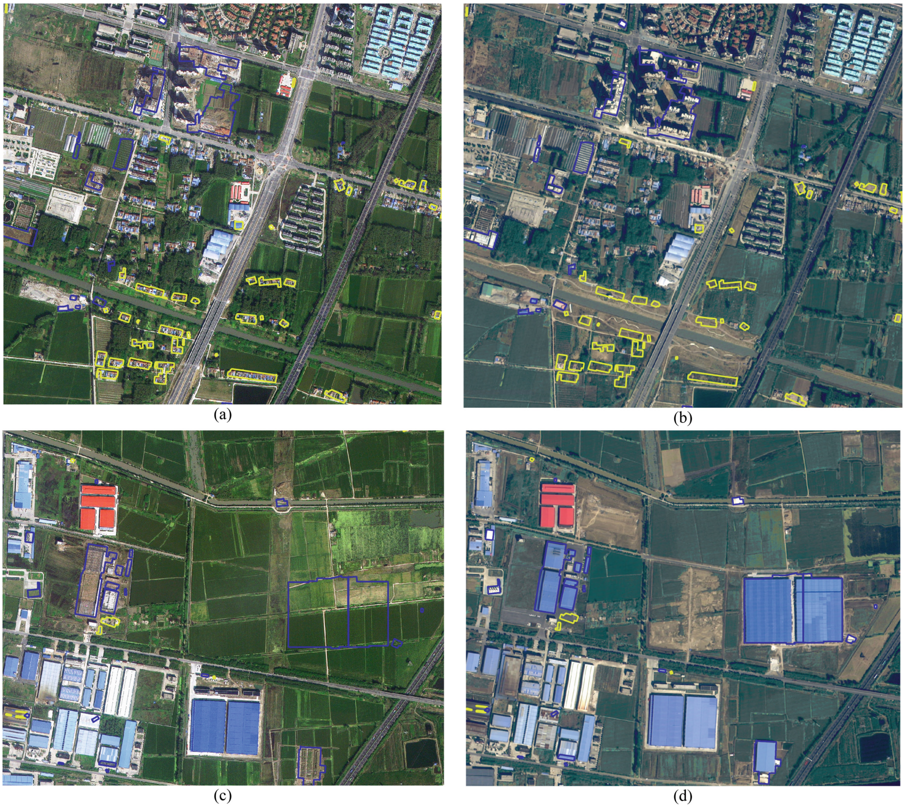

Fig. 4. Detailed building semantic change comparison. (The blue color and yellow color are building increase, building decrease, respectively). (a) and (c) is the former time-phase image, (b) and (d) is the latter time-phase image.

Table 2. Accuracy assessment

Accuracy	Recall	F1-score
0.9219	0.8193	0.8094

We use ArcGIS10.6 to statistic the area of building increase and decrease, the result is as follows (Table 3):

Table 3. Building increase and decrease area

Building increase (km^2)	Building decrease (km^2)
1.528	0.3987

We can see that the area of building increase is relatively faster than the decrease, one main reason is as the economic development of Jiangsu Province, the area of residential and industrial also increased.

2.6 Discussion

Though our deep learning model can detect building change in very high-resolution images, the detection accuracy on other median resolution images may be unsatisfactory, for example in 10m spatial resolution images. Another limitation of our method is it can detect building change with only RGB images, while in multispectral images, such as with an auxiliary near-red band, it cannot use the full spectral information, thus our next step is to use the multispectral band to train our multi-task deep learning model.

In order to help other researchers to capture building semantic change quickly, we also designed a building change detection software, which is user-friendly and high-performance. The graphical user interface is as follows (Fig. 5):

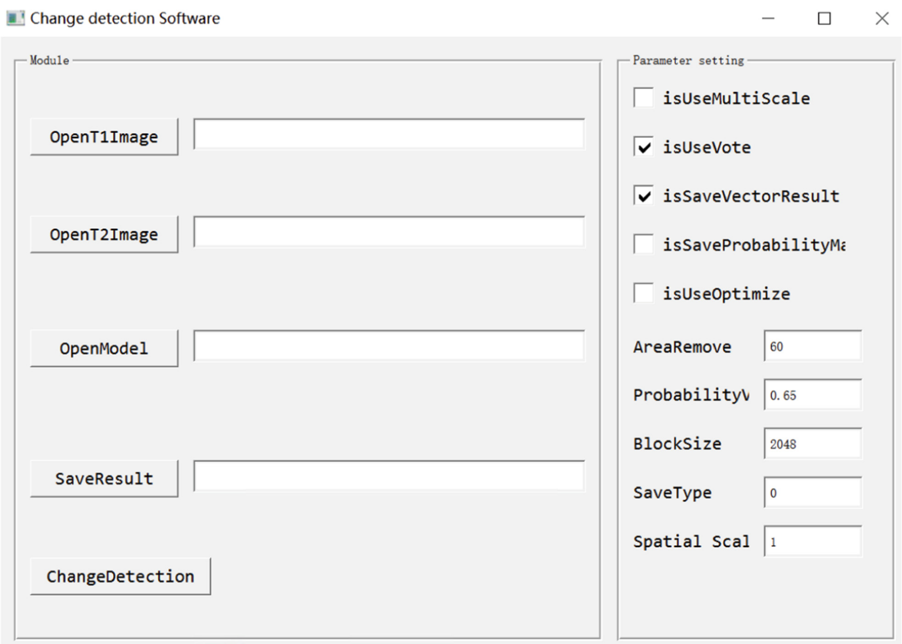

Fig. 5. Building change detection software

If users want to use this software freely, please contact the author of this manuscript. This software was developed on CUDA11.3, thus, on a client computer, users must install CUDA11.3 to get high-performance computing.

References

1. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. Nature **521**(7553), 436–444 (2015)
2. Liu, Y., Pang, C., Zhan, Z., et al.: Building change detection for remote sensing images using a dual-task constrained deep siamese convolutional network model. IEEE Geosci. Remote Sens. Lett. **18**(5), 811–815 (2020)
3. Shen, Q., Huang, J., Wang, M., et al.: Semantic feature-constrained multitask siamese network for building change detection in high-spatial-resolution remote sensing imagery. ISPRS J. Photogramm. Remote. Sens. **189**, 78–94 (2022)
4. Ji, S., Shen, Y., Lu, M., et al.: Building instance change detection from large-scale aerial images using convolutional neural networks and simulated samples. Remote Sens. **11**(11), 1343 (2019)
5. Dong X, Qian L, Guan Y, et al. A multiclass classification method based on deep learning for named entity recognition in electronic medical records. In: 2016 New York Scientific Data Summit (NYSDS), pp. 1–10. IEEE (2016)
6. Zhu, Y., Newsam, S.: Densenet for dense flow. In: 2017 IEEE International Conference on Image Processing (ICIP), pp. 790–794. IEEE (2017)
7. Szegedy, C., Ioffe, S., Vanhoucke, V., et al.: Inception-v4, inception-resnet and the impact of residual connections on learning. Thirty-first AAAI Conference on Artificial Intelligence (2017)
8. Pant, G., Yadav, D.P., Gaur, A.: ResNeXt convolution neural network topology-based deep learning model for identification and classification of Pediastrum. Algal Res. **48**, 101932 (2020)
9. Ge, Z., Liu, S., Wang, F., et al.: Yolox: exceeding yolo series in 2021. arXiv preprint arXiv:2107.08430 (2021)
10. Botev, Z.I., Kroese, D.P., Rubinstein, R.Y., et al.: The cross-entropy method for optimization. In: Handbook of Statistics. vol. 31, pp. 35-59. Elsevier (2013)
11. Garcia-Garcia, A., Orts-Escolano, S., Oprea, S., et al.: A review on deep learning techniques applied to semantic segmentation. arXiv preprintarXiv:1704.06857 (2017)
12. Aurelio, Y.S., de Almeida, G.M., de Castro, C.L., et al.: Learning from imbalanced data sets with weighted cross-entropy function. Neural Process. Lett. **50**(2), 1937–1949 (2019)
13. Lin, T.Y., Goyal, P., Girshick, R., et al.: Focal loss for dense object detection. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 2980–2988 (2017)
14. Shi, F., Qian, H., Chen, W., et al.: A fire monitoring and alarm system based on YOLOv3 with OHEM. In: 2020 39th Chinese Control Conference (CCC), pp. 7322–7327. IEEE (2020)
15. Diakogiannis, F.I., Waldner, F., Caccetta, P., et al.: ResUNet-a: a deep learning framework for semantic segmentation of remotely sensed data. ISPRS J. Photogramm. Remote. Sens. **162**, 94–114 (2020)