Detecting Building boundary Using Residual U-Net

Suhong Yoo (1), Jisang Lee (1), Junsu Bae (1), Honggyoo Sohn (1)

[1] Yonsei Univ., 50, Yonsei-ro, Seodaemun-gu, Seoul, Republic of Korea
Email: swennoir@yonsei.ac.kr; ontheground@yonsei.ac.kr;
junsu510@yonsei.ac.kr; sohn1@yonsei.ac.kr

KEY WORDS: Building detection, Residual U-Net, Segmentation

ABSTRACT: In order to automatically extract buildings from satellite images, we conducted experiments using Residual U-Net, one of semantic segmentation algorithms. Although limitations have been found, it has the advantage of automatically and quickly acquiring the building information which has been manually extracted by humans in the past.

1. Introduction

Research and development are underway to use deep-learning algorithms in different fields due to the development and distribution of several open libraries such as Tensorflow, Pytorch, Caffe etc. and the improvement of hardware performance. Several challenging competitions such as the ImageNet Challenge have been shown dramatically improved in the accuracy of image-based classification, which further accelerates the researchers' application in deep-learning techniques to their research areas. Especially, in the field of remote sensing, there are many data that can be produced in the form of image or 3D voxel which can be easily used in the deep-learning algorithm, and various studies using the deep learning algorithm have been announced recently.

One of the semantic segmentation algorithms, U-Net, is an algorithm published in the field of medical imaging in 2015 and used to extract the boundaries of cells photographed by electron microscopy. If we use max pooling in the neural network, the location information is missing, and if we do deconvolution again, there is a possibility of learning in an undesired direction. U-net is a meaningful algorithm because it improves the learning accuracy of deconvolution, which is known to be difficult to learn by using concatenation. This is a great advantage in applying remote sensing data. This is because the disturbance for learning a large-sized image is reflected in the same way as the high-resolution remote sensing data.

In this study, using the residual u-net improved U-net, experiments were carried out to extract high-precision building masks by adjusting hyperparameters based on previous studies.

2. Methods and results

The automatic building detection was performed using Residual U-Net which was an improved U-Net and an accuracy evaluation was performed. The most important thing in deep learning is to have reliable data. In the 2D Semantic Labeling Contest in ISPRS, the RGB, NIR and Digital Surface Model (DSM) are released in some parts of Potsdam, Germany. The data were processed in a specific way that is easy to use for learning. Starting with a learning rate of 1e-5, the learning rate is reduced for every 100 epochs for a total of 500 epochs and so the optimal solution can be learned. We used various loss functions and pooling functions for improving the precision. As a result of learning, the model using all the data and the building mask, it was found to have 99% of precision when the binary cross entropy and max pooling function were adopted.

3. Conclusion

In this study, Residual U-Net was employed, which is used in many studies to automatically detect buildings. Because of the different shapes of the building and the color of the roof, there is still a

limit to the detection and extraction of the perfect shape of the building. Compared to previous methods of deep-learning techniques, we could extract a considerable number of buildings and similar shapes.

On the other hand, there are some features that are not understood by ground truth or that are unclear enough to be mistaken as buildings even when judged by humans, so further research is needed to address this issue.

4. Acknowledge

This research was supported by a grant (2019-MOIS32-015) of Disaster-Safety Industry Promotion Program funded by Ministry of Interior and Safety (MOIS, Korea).

5. References

Boonpook, W.; Tan, Y.; Ye, Y.; Torteeka, P.; Torsri, K.; Dong, S., 2018. A Deep Learning Approach on Building Detection from Unmanned Aerial Vehicle-Based Images in Riverbank Monitoring. Sensors, 18, 3921.

ISPRS 2D Semantic Labeling Contest, Retrieved 30, August 30, 2019, from http://www2.isprs.org/commissions/comm3/wg4/semantic-labeling.html

Zhang, Z.; Liu, Q.; Wang, Y., 2018. Road extraction by deep residual u-net. IEEE Geosci. Remote Sens. Lett.