International Cartographic Association
Association Cartographique Internationale

Investigating the Importance of Digital Surface Models for Building Segmentation with Deep Learning

Güren Tan Dinga [a]

[a] *Lab for Geoinformation and Geovisualization (g2lab), HafenCity University Hamburg, Germany, gueren.dinga@hcu-hamburg.de*

Keywords: Deep Learning, Building Segmentation, Elevation Data, Digital Surface Models

Abstract:

Image segmentation is a fundamental task in computer vision and finds application in many other disciplines, such as surveying, medical image analysis and augmented reality. The main idea behind image segmentation is to partition an image into multiple meaningful regions or segments of pixels that are semantically connected. The concept behind image segmentation dates back to the early days of computer vision research, as being mentioned by Udupa and Samarasekera (1996). Since then, various segmentation algorithms and methods have been developed and proposed. While segmentation tasks were conducted with traditional methods such as thresholding or k-means clustering, advancements over the last decade focus on deep learning with convolutional neural networks (CNNs).

One of the most influential and impactful CNNs for segmentation tasks was the U-Net, proposed by Ronneberger et al. (2015). The encoder-decoder network was able to surpass the state-of-the-art for medical tasks at the time and was shortly thereafter used for segmentation tasks in numerous other disciplines, including geoinformatics. More recent networks, for example the Mask R-CNN (He et al. (2017)) or DeepLabV3+ (Chen et al. (2018)), could show that it is possible to learn even more intricate patterns and features from images to achieve state-of-the-art performances in segmentation tasks regardless of the domain (e.g. terrestrial scene segmentation, medical image segmentation or satellite image segmentation).

Consequently, a range of publications could demonstrate that building segmentation with the help of deep learning is possible with high accuracy. However, for various segmentation tasks the knowledge about how neural networks make decisions are still lacking.

My ongoing research combines domain-specific knowledge with a multi-modal approach. It utilizes not only two-dimensional (images consisting of red, green, blue and infrared bands) but also three-dimensional information (digital surface models or DSMs) to segment aerial images. The preliminary results are in accordance with publications investigating the impacts of elevation information for building segmentation tasks, and show a higher accuracy than models that have been trained using only two-dimensional data. To highlight the difference between results of models that were trained with (1) DSMs only, (2) RGB data, (3) RGBI data, (4) RGBI and DSMs and (5) RGB and DSMs, the deep learning models were trained using different data arrangements. This means that in some experiments (2) and (3), elevation data were left out to emphasize their influence on the final segmentation results. Additionally, experiment (1) was conducted using only the elevation data to highlight the influence of two-dimensional data when comparing the corresponding results. Figure 1 shows a diagram of the simplified workflow.

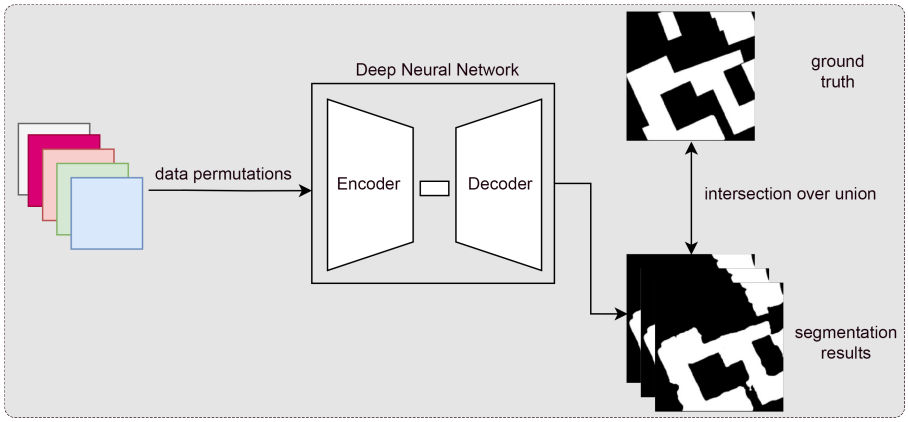

Figure 1. Simplified diagram to show the workflow for different arrangements of the blue, green, red, infrared and DSM bands. After training the neural network with the respective arrangement, the results are evaluated using, among other metrics, the intersection over union.

https://doi.org/10.5194/ica-abs-7-30-2024

Preliminary results indicate that edge cases, in particular, benefit from using various levels of information available. While solitary buildings without any occlusion were segmented with high accuracy independent of the amount of information used during the training phase, elevation data seems to play a significant role in the segmentation results of buildings with green roofs and contribute to more accurate segmentation outcomes. Conversely, we have identified cases in which the addition of elevation data led to worse results. The findings from investigating the segmentation results where adding more data leads to a lower performance seem to be systematic and need to be quantified. To avoid network- and data-dependent influences on our segmentation results, we have trained two deep learning models of different depth (U-Net and DeepLabV3) using two datasets (a custom dataset created with open data from North Rhine-Westphalia and the ISPRS Potsdam dataset).

In the current phase of our research project, we are using occlusion and perturbation-based methods to investigate the influence of different parts of our input data (note that the occlusion-based methods here are different from occlusions in the image itself caused by e.g. vegetation covering a part of a building. The occlusions and perturbations are controlled by the user and cover predefined or randomly sampled parts of the image). The presentation will provide an overview of the current state of the ongoing research, have an emphasis on the challenges, and show findings as well as outline future work.

References

Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F. and Adam, H., 2018. Encoder-decoder with atrous separable convolution for semantic image segmentation. In: *Proceedings of the European Conference on Computer Vision (ECCV)*.

He, K., Gkioxari, G., Dollár, P. and Girshick, R., 2017. Mask r-cnn. In: *2017 IEEE International Conference on Computer Vision (ICCV)*, pp. 2980–2988.

Ronneberger, O., Fischer, P. and Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation. In: N. Navab, J. Hornegger, W. M. Wells and A. F. Frangi (eds), *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, Springer International Publishing, Cham, pp. 234–241.

Udupa, J. K. and Samarasekera, S., 1996. Fuzzy connectedness and object definition: Theory, algorithms, and applications in image segmentation. *Graphical Models and Image Processing* 58(3), pp. 246–261.

https://doi.org/10.5194/ica-abs-7-30-2024