International Cartographic Association
Association Cartographique Internationale

Building Footprint Extraction from High Resolution UAV Images Using Deep Learning Algorithms in the Context of Unplanned Urbanisation

Halil İbrahim Şenol[a,*], Arzu Çöltekin[b]

[a] *Harran University, halilisenola@gmail.com*
[b] *Institute of Interactive Technologies,University of Applied Sciences and Arts Northwestern Switzerland, arzu.coltekin@fhnw.ch*

* Corresponding author

Keywords: building extraction, deep learning, 3D city modelling, UAV

Abstract:

With the acceleration of urbanisation, we also see a strong increase in *unplanned urbanisation* globally. Especially in underdeveloped and developing countries, unplanned urbanisation is an important reality, and causes serious problems e.g., in the quality of urban living, infrastructure and health services, access to education *etc*. In this context, building extraction from imagery is an important application, because it is frequently used for urban planning and change detection, e.g., in monitoring urban developments and accessibility, or monitoring street-level changes after natural disasters (Boonpook et al., 2021). For extracting building footprints, high spatial (and in the case of disaster, also temporal) resolution images, such as the one shown in Figure 1, are needed, so that building details and sizes can be identified with high precision. This can be important in subsequent decision making, e.g., for interventions, required services or building health. Although various manual and computational methods are used for building extraction, due to the promising speed and precision they offer, use of machine learning and deep learning algorithms has been increasing in recent years (e.g., Tiede et al., 2021). On the other hand, unplanned urbanisation can cause challenges to deep learning due to irregularities in spatial structures.

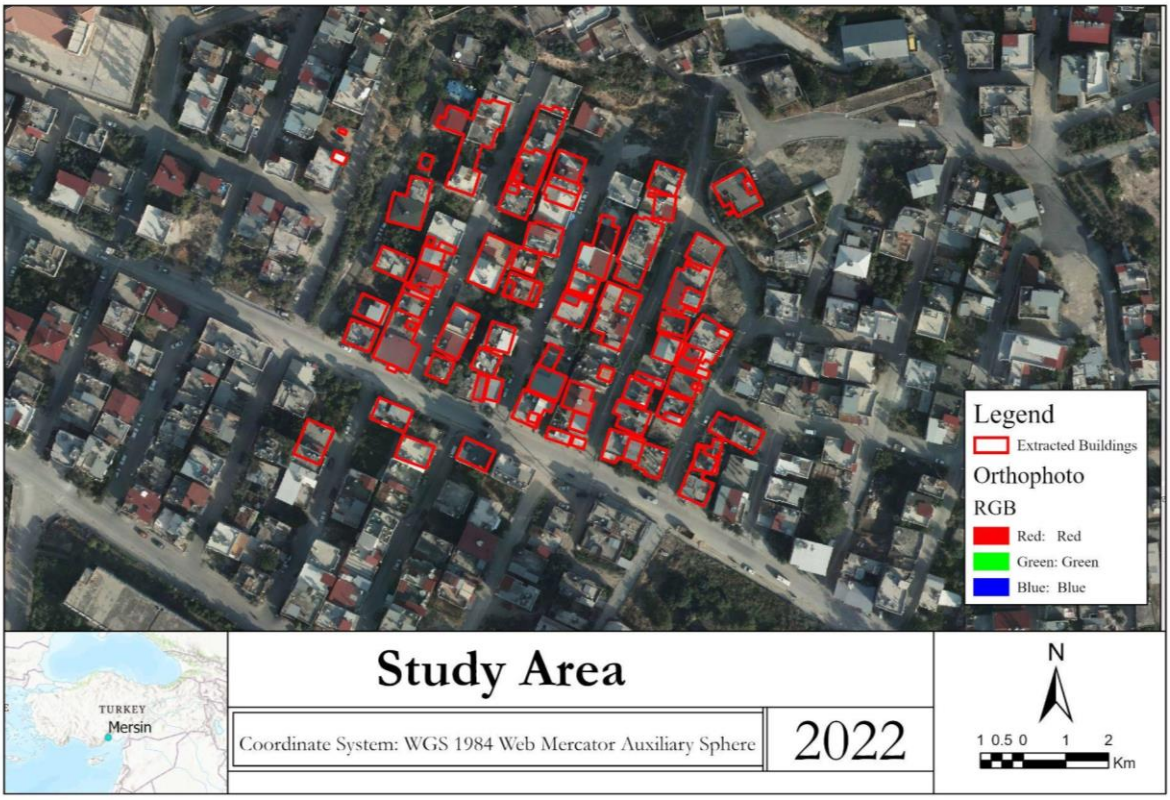

Figure 1. Study area. The image is obtained in a UAV mission conducted by the authors.

In this study, we apply a deep learning approach to extract building footprints in a region with unplanned urbanisation, i.e., we use orthophotos created based on 10 cm resolution unmanned aerial vehicle (UAV) photographs of Mersin province in Turkey. Specifically, we used Sensefly Ebee Plus UAV with a built-in 20 MP camera for photogrammetric data collection. Because of its mixed content for planned and unplanned urbanisation, a neighbourhood in the centre of Mersin was selected for the study. In the application, due to its previously successful use in object detection (Wu et al., 2020), we used the Mask R-CNN deep learning algorithm (He et al., 2017) and carried the application out through ArcGIS Pro (URL-1). Mask R-CNN algorithm is a tool used for spatial image segmentation and object recognition (Su et al., 2019), and because it is the most compatible algorithm with the Deep Learning tools in ArcGIS Pro, the project's original software environment. The approach automatically detects objects with specific shapes on the high-resolution image, while also creating a high-quality segmentation mask for each sample. We used a pre-trained package for ArcGIS Pro Deep Learning tool on a very high resolution orthophoto data set which was obtained through UAV flights that we planned and conducted ourselves. Besides its exploratory nature with the areas that contain structures from unplanned urbanisation, an additional novelty in this research is to examine if automatically extracted building footprints can be used in 3D GIS and open source data platforms.

The algorithm identified the target buildings with an accuracy of 79% and extracted their footprints (Figure 1). While 79% accuracy is fairly high, it is not as good as many modern applications of object recognition with machine / deep learning algorithms. The key explanation here is possibly that the model is trained with a disjoint dataset (where there are gaps between each house), whereas in these examples the data contains twin-buildings without gaps between them based on an adjoining system applied in unplanned urban areas throughout Turkey. Unsurprisingly, thus, the algorithm perceived some of the building groups as single buildings and could not distinguish individual buildings. Possible reasons are the irregularity of the spatial structures we mentioned earlier, i.e., the proximity of buildings in unplanned urban areas, the presence of different objects in the areas belonging to the building, and the insufficient width of the roads between the buildings bring special challenges for the application of building extraction in unplanned urbanisation. Consequently, to facilitate the use of the extracted buildings in urban planning applications, we create solid models of buildings of known height (not included in this submission). Our experience suggests that not only this deep-learning approach allows useful building footprint extraction, but also the consequent 3D models can be used for visualisation purposes in smart city applications. Furthermore, building footprints produced from high-resolution orthophotos can be used to enrich data in areas with poor OpenStreetMap (OSM) data, thus contributing towards creating cartographic content where ground surveys are not always possible. A useful next step would be to train the deep learning models with more versatile data to increase the accuracy of the outcomes even further.

References

Boonpook, W., Tan, Y., & Xu, B. (2021). Deep learning-based multi-feature semantic segmentation in building extraction from images of UAV photogrammetry. *International Journal of Remote Sensing*, 42(1), 1-19.

Tiede, D., Schwendemann, G., Alobaidi, A., Wendt, L., & Lang, S. (2021). Mask R-CNN-based building extraction from VHR satellite data in operational humanitarian action: An example related to Covid-19 response in Khartoum, Sudan. *Transactions in GIS*, 25(3), 1213-1227.

Wu, M., Yue, H., Wang, J., Huang, Y., Liu, M., Jiang, Y., Ke, C., & Zeng, C. (2020). Object detection based on RGC mask R-CNN. *IET Image Processing*, 14(8), 1502-1508.

He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask r-cnn. *In Proceedings of the IEEE international conference on computer vision* (pp. 2961-2969).

URL-1: https://www.arcgis.com/home/item.html?id=4e38dec1577b4b7da5365294d8a66534

Su, H., Wei, S., Yan, M., Wang, C., Shi, J., & Zhang, X. (2019). Object detection and instance segmentation in remote sensing imagery based on precise mask R-CNN. *In IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium* (pp. 1454-1457).

Abstracts of the International Cartographic Association, 5, 2022.
European Cartographic Conference – EuroCarto 2022, 19–21 September 2022, TU Wien, Vienna, Austria.
https://doi.org/10.5194/ica-abs-5-144-2022