Investigating the capability of UAV imagery for AI-assisted mapping of Refugee Camps in East Africa

Christopher Yan-Chak Chan[1,*], Matthias Weigand[2], Emran Alchikh Alnajar[3], Hannes Taubenböck[1,2]

[1] Department of Remote Sensing, Julius Maximilians University of Würzburg, Würzburg, Germany; chrischank@protonmail.ch, yan_chak_christopher.chan@stud-mail.uni-wuerzburg.de
[2] German Remote Sensing Data Center (DFD), German Aerospace Center (DLR), Weßling, Germany; matthias.weigand@dlr.de, hannes.taubenboeck@dlr.de
[3] Humanitarian OpenStreetMap Team, Washington D.C., USA; omran.najjar@hotosm.org

* Author to whom correspondence should be addressed.
This abstract was accepted to the Academic Track of the State of the Map 2022 Conference after peer-review.

Refugee camps and informal settlements provide accommodation to some of the most vulnerable populations, with many of them located in Sub-Saharan Africa. Many of these settlements lack up-to-date geoinformation that we take for granted in developed settlements. Having up-to-date maps on the dimension, spatial layout are important for assisting administration tasks such as crisis intervention, infrastructure development, and population estimates which encourage economic productivity [1]. The data inequality between the developed and developing countries are results of multiple reasons ranging from a lack of commercial interest to knowledge gaps in data contributors [2–4]. Such disparity can be reduced using assisted mapping technology. A combination of Very High Resolution (VHR) satellite imagery and Machine Learning (ML) based methods that exploit the textural, spectral, and morphological characteristics of VHR imagery are commonly used to extract geospatial and imagery of these complex environments [5]. Although many have shown promising results in satellite VHR scenarios (e.g. differentiating slum and non-slum [6,7]), in drone imagery, however, results might suffer from noise caused by environment and drone-based specific motion problems. Recent advances in Deep Learning (DL) based Computer Vision (CV), however, might be able to address these issues [8].

In our pilot project, we investigated the capabilities of applying DL semantic segmentation methods for delineating building footprints in refugee camps from open-data drone imagery. This study is connected to a larger initiative to open-source the AI assisted mapping platform in the current Humanitarian OpenStreetMap Team's (HOT) ecosystem. The study focuses on two refugee camps in East Africa located in a similar savannah ecosystem. The first camp is located in Dzaleka, Dowa, Malawi, this area is split into the Dzaleka North and Dzaleka main camp, home to ca. 40,000 refugees. The northern camp is characterised by newer, spatially well-planned metal-sheeted roofs, while the southern main

Chan, C.Y., Weigand, M., Alnajar, E.A., & Taubenböck, H. (2022). Investigating the capability of UAV imagery in AI-assisted mapping of Refugee Camps in East Africa.
In: Minghini, M., Liu, P., Li, H., Grinberger, A.Y., & Juhász, L. (Eds.). Proceedings of the Academic Track at State of the Map 2022, Florence, Italy, 19-21 August 2022. Available at https://zenodo.org/communities/sotm-22
DOI: 10.5281/zenodo.7004576

camp is characterised by arrangements of dense mud-walled buildings with stone-lined thatched-roofs [9]. The second camp, the Kalobeyei settlement is home to ca. 34,800 refugees as of 2019, the settlement part of an extension for the larger Kakuma refugee camp, located in the rural county of Turkana, North-West Kenya. This camp is significantly more spacious and is characterised by spatially well-planned shelters with metal-sheeted roofs [10]. VHR drone images were provided for both camps by the OpenAerialMap project and vector labels produced by HOT volunteers were provided for the Dzaleka camps, and vector labels for the Kalobeyei settlements were created specifically for this study.

DL and semantic segmentation applications aim at classifying each pixel of an image into a predefined set of semantic classes. Semantic segmentation image classification tasks require high quality of reference data. In particular, the motion artefact problem [8] needs to be addressed. However, in the case of this study, a large quantity of available labels from the Dzaleka camps were not created with these challenges in mind. These imperfections in labelling could cause the trained model to misclassify. In order to have a model which performs well on drone imagery, we hypothesise that this will be a significant feature for the model to learn. Therefore, the Kalobeyei dataset was labelled to be the pixel-aligned dataset accounting for motion artefacts. Through this unique combination of different data quality levels, their influence on the classification results could be tested.

Several architectures of Convolutional Neural Networks (herein CNN) have recently developed. From these, the U-Net architecture [11] was selected on the basis of applicable ability in many domains, a proven track record of performance in remote sensing segmentation tasks, and relative computational efficiency [12–14]. The symmetrical encoder-decoder type architecture is able to extract deeper features in the encoder layers, while the decoder layers recover and interpolate spatial features [15]. The ability to switch out the encoder structure allows the DL practitioner to experiment with more up-to-date architectures without changing the output shape. This drastically increases the combination of experiments that allow testing the best combinations of encoder-decoder structure suitable for the dataset. All the experiments in this study were carried out using the high-level PyTorch API Segmentation-Model-PyTorch [16]. In this study, the experiments with changed encoder came from the EfficientNet family. There are three reasons for this selection: Firstly, at one of the last stages of the Open-Cities-AI-Challenge (OCC) competition winning network [17], EfficientNet B1 was used as an encoder. Secondly, the EfficientNet family is a set of network architectures that are structured and easy to scale up when computational resources become available. Thirdly, they are an accepted representation of generalised state-of-the-art architectures that have been tested and performed well in classical CV datasets [18]. In essence, these are sets of experiments that mix and match old and new architectural design. Regarding data pre-processing, the drone imagery was normalised and resampled to 15 cm resolution, then subsequently cropped at two-thirds overlapping steps to increase the quantity of training data. The training, validation, and testing dataset were split on a 60, 30, and 10% ratio. With appropriate augmentation techniques applied to the training and validation dataset.

The objective of this study is to test the U-Net and several variations of the architectures' performance for building footprint mapping, initially on the pixel-aligned and simple Kalobeyei dataset. Subsequently, the models were introduced to the less-aligned Dzaleka and Dzaleka North datasets of higher complexity. A comparison of the baseline experiments, which kept the settings of hyperparameters to be consistent, were conducted between the architectures. Class-based accuracy assessments metrics were used to evaluate performances between the baseline experiments setup. This allows evaluation into which level of data quality is required to achieve acceptable classification results when scaling DL assisted mapping efforts in the context of humanitarian mapping.

Initial baseline experiments suggest limited transferability from the competition winning OCC model. This indicates that the OCC model might be over-generalised to the competition test dataset, which mainly consists of a similar drone imagery of 10 different urban areas in Africa. This is accentuated by this model achieving very high confidence on metal-sheeted roofs, while not detecting any of the more complicated thatched roofs common in the Dzaleka camp (see Figure 1, left). The OCC model was found to also struggle with some of the more obscure drone motion artefacts occurring at the edge of the imagery in the Kalobeyei camp.

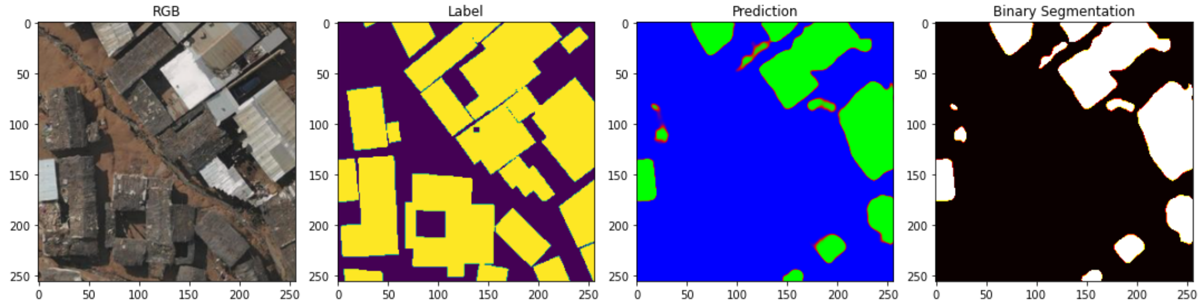

Figure 1. Prediction example from further training of EfficientNet B1 U-Net of the OCC winning network.

In most of our experiments, precision and recall measures have both reached above 0.7. However, there are still significant variations among different architectures and between the datasets. Precision and recall suggest that unmodified U-Nets were least affected by the introduction of the less accurate samples from the Dzaleka dataset. Meanwhile, in absolute terms, the EfficientNet encoder U-Nets performed better if only the Kalobeyei dataset is considered. The results show that deeper versions of the respective network architecture had not universally produced improvements, but they vary by dataset. Segmentation results from the OCC model (see Figure 1), when adapted to the settlements of this study, did not outperform the other tested architectures.

This study demonstrated the ability to use DL semantic segmentation to perform building segmentation in complex humanitarian applications. Having increased access to open-data VHR drone imagery such as the OpenAerialMap initiative is an advantage to building AI-assisted humanitarian mapping. The study evaluated various U-Net based architectures and data input setup. Yet, the variation of the results not only emphasised the complexity of DL based methods, but also indicate that further efforts will be needed to focus on the selection of suitable network architectures as they show varying resilience towards imperfect reference samples.

References

[1] Herfort, B., Lautenbach, S., de Albuquerque, J.P., Anderson, J., & Zipf, A. (2021). The evolution of humanitarian mapping within the OpenStreetMap community. *Scientific Reports, 11*(1), 1–15.

[2] Anderson, J., Sarkar, D., & Palen, L. (2019). Corporate Editors in the Evolving Landscape of OpenStreetMap. *ISPRS International Journal of Geo-Information, 8*(5), 232.

[3] Veselovsky, V., Sarkar, D., Anderson, J., & Soden, R. (2021). An automated approach to identifying corporate editing activity in OpenStreetMap. In: Minghini, M., Ludwig, C., Anderson, J., Mooney, P., & Grinberger, A.Y. (Eds.) *Proceedings of the Academic Track at the State of the Map 2021 Online Conference*, 31–33.

[4] Yang, A., Fan, H., Jing, N., Sun, Y., & Zipf, A. (2016). Temporal Analysis on Contribution Inequality in OpenStreetMap: A Comparative Study for Four Countries. *ISPRS International Journal of Geo-Information, 5*(1), 5.

[5] Taubenböck, H., Kraff, N.J., & Wurm, M. (2018). The morphology of the Arrival City—A global categorization based on literature surveys and remotely sensed data. *Applied Geography, 92*, 150–167.

[6] Kuffer, M., Pfeffer, K., Sliuzas, R., & Baud, I. (2016). Extraction of Slum Areas From VHR Imagery Using GLCM Variance. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 9*(5), 1830–1840.

[7] Wurm, M., Weigand, M., Schmitt, A., Geiß, C., & Taubenböck, H. (2017). Exploitation of textural and morphological image features in Sentinel-2A data for slum mapping. *2017 Joint Urban Remote Sensing Event (JURSE)*, 1–4.

[8] Smith, M.W., Carrivick, J.L., & Quincey, D.J. (2016). Structure from motion photogrammetry in physical geography. *Progress in Physical Geography: Earth and Environment, 40*(2), 247–275.

[9] UNHCR (2014). Malawi: Joint Assessment Mission Report. Dzaleka Refugee Camp (November 2014). Retrieved from https://www.unhcr.org/protection/health/5680f7d09/malawi-joint-assessment-mission-report-dzaleka-refugee-camp-november-2014.html

[10] UNHCR & Ministry of Foreign Affairs of Denmark (2019). UNHCR/Danida - The Integrated Solutions Model in and Around Kalobeyei, Turkana, Kenya. Retrieved from https://www.unhcr.org/research/evalreports/5dfa287c4/unhcrdanida-integrated-solutions-model-around-kalobeyei-turkana-kenya.html

[11] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab, N., Hornegger, J., Wells, W.M., & Frangi, A.F. (Eds.) *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015,* Springer, Cham, 234–241.

[12] Hoeser, T., Bachofer, F., & Kuenzer, C. (2020). Object Detection and Image Segmentation with Deep Learning on Earth Observation Data: A Review—Part II: Applications. *Remote Sensing, 12*(18), 3053.

[13] Hoeser, T., & Kuenzer, C. (2020). Object Detection and Image Segmentation with Deep Learning on Earth Observation Data: A Review-Part I: Evolution and Recent Trends. *Remote Sensing, 12*(10), 1667.

[14] Xu, Y., Wu, L., Xie, Z., & Chen, Z. (2018). Building Extraction in Very High Resolution Remote Sensing Imagery Using Deep Learning and Guided Filters. *Remote Sensing, 10*(1), 144.

[15] Wurm, M., Stark, T., Zhu, X.X., Weigand, M., & Taubenböck, H. (2019). Semantic segmentation of slums in satellite images using transfer learning on fully convolutional neural networks. *ISPRS Journal of Photogrammetry and Remote Sensing, 150*, 59–69.

[16] Yakubovyskiy, P. (2021). Segmentation Models documentation. Retrieved from https://segmentation-models-pytorch.readthedocs.io/en/latest/index.html

[17] Open Cities AI Challenge: Segmenting Buildings for Disaster Resilience. (2022). Retrieved from https://github.com/drivendataorg/open-cities-ai-challenge

[18] Tan, M., & Le, Q.V. (2020). EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. *arXiv preprint*, arXiv:1905.11946.