

EGU2020-1242

<https://doi.org/10.5194/egusphere-egu2020-1242>

EGU General Assembly 2020

© Author(s) 2021. This work is distributed under the Creative Commons Attribution 4.0 License.



## Sub-pixel classification of anthropogenic features using deep-learning on Sentinel-2 data

Nikolaos Ioannis Bountos<sup>1,2</sup>, Melanie Brandmeier<sup>1</sup>, and Mark Günter<sup>3</sup>

<sup>1</sup>Esri Germany and Switzerland

<sup>2</sup>Technical University Munich

<sup>3</sup>University of Applied Sciences Bern

Urban landscapes are characterized as the fastest changing areas on the planet. However, regularly monitoring of larger areas it is not feasible using UAVs or costly air borne data. In these situations, satellite data with a high temporal resolution and large field of view are more appropriate but suffer from the lower spatial resolution (deca-meters). In the present study we show that by using freely available Sentinel-2 data from the Copernicus program, we can extract anthropogenic features such as roads, railways and building footprints that are partly or completely on a sub-pixel level in this kind of data. Additionally, we propose a new metric for the evaluation of our methods on the sub-pixel objects. This metric measures the performance of the detection of an object while penalizing the false positive classification. Given that our training samples contain one class, we define two thresholds that represent the lower bound of accuracy for the object to be classified and the background. We thus avoid a good score in occasions where we classify correctly our object, but a wide area of the background has been included in our prediction. We investigate the performance of different deep-learning architectures for sub-pixel classification of the different infrastructure elements based on Sentinel-2 multispectral data and the labels derived from the UAV data. Our study area is located in the Rhone valley in Switzerland where very high-resolution UAV data was available from the University of Applied Sciences. Highly accurate labels for the respective classes were digitized in ArcGIS Pro and used as ground-truth for the Sentinel data. We trained different deep learning models based on state-of-the-art architectures for semantic segmentation, such as DeepLab and U-Net. Our approach focuses on the exploitation of the multi spectral information to increase the performance of the RGB channels. For that purpose, we make use of the NIR and SWIR 10m and 20m bands of the Sentinel-2 data. We investigate early and late fusion approaches and the behavior and contribution of each multi spectral band to improve the performance in comparison to only using the RGB channels. In the early fusion approach, we stack nine (RGB, NIR, SWIR) Sentinel-2 bands together, pass them from two convolutions followed by batch normalization and relu layers and then feed the tiles to DeepLab. In the late fusion approach, we create a CNN with two branches with the first branch processing the RGB channels and the second branch the NIR/SWIR bands. We use modified DeepLab layers for the two branches and then concatenate the outputs into a total output of 512 feature maps. We then reduce the dimensionality of the result into the final output equal to the number of classes. The dimension reduction step happens in two convolution layers.

We experiment on different settings for all of the mentioned architectures. In the best-case scenario, we achieve 89% overall accuracy. Moreover, we measure 60% building accuracy, streets accuracy 60%, railway accuracy 73%, river accuracy 92% and background accuracy 94%.