# Extended Abstract: Deep learning-based road segmentation using aerial imagery for automated change detection

Course: Integrated Project: Interdisciplinary | Integrated | Interactive

Author: Adian Dawuda (adian.dawuda@stud.sbg.ac.at | s1093093)

Delivery date: 31.03.2023

Version: 1.0

## Introduction

The field of earth observation (EO) provides a growing amount of earth data, including satellite and aerial imagery (Gomes et al., 2020; Sudmanns et al., 2020). EO data may be used to monitor the earth's surface and identify potential changes for numerous application domains such as urbanization and planning, disaster management, agriculture, etc. Sudmanns et al. (2020) also describe the need for EO data to be interpreted to gather meaningful information from it. For aerial or satellite images, this traditionally includes the separate steps of segmentation and classification (Lang et al., 2019; Matsuyama, 1987). Through research and growingly accessible computing power, deep learning has emerged as a promising approach to extracting information from images and can combine segmentation and classification in one step referred to as semantic segmentation (H. Wang et al., 2017; Wu et al., 2019). This project aims to perform a change detection analysis on specific sections of the road network in Cologne, Germany using a U-Net architecture-based convolutional neural network (CNN), first implemented by Ronneberger et al. (2015), to generate binary semantic segmentation masks of roads from aerial images. Automating such manually laborious tasks has great potential to save time and improve work efficiency.

As there is no labeled data available for the area of interest, the project uses the road dataset of the *Massachusetts Road and Building Detection Datasets* (Mnih, 2013) for training. This approach will test the model's ability to generalize to new data. The training labels for Massachusetts roads are based on OpenStreetMap data, while the test labels for Cologne roads are created based on manual delineation.

Two important factors shall be considered to assess the quality of the model for road change detection. The first is the accuracy of the segmentation, meaning how well the predicted target maps, represent the ground truth target maps. The second factor is the time and effort it takes to conduct the analysis with the developed model compared to a manual approach. further factors such as cost, accessibility, and transparency may also play a role. However, these are largely outside of the scope of this paper and will not be discussed in detail. Given the interdisciplinary application domains of image understanding and change detection, this project may provide valuable findings for a broad range of actors and contributes to the growing field of earth observation and deep learning. The processes and findings of this project are outlined in this paper. All code and the pre-trained model are available on GitHub and GitLab.

# Material and Methods

The creation of the U-net CNN model and the necessary data preprocessing and postprocessing steps are conducted using the Python programming language. The TensorFlow framework is used to build, use, and evaluate the model. The code will be deployed in a Jupyter notebook, hosted on Google Collab. This grants access to a powerful computing environment and allows for easily sharable code and reproducible results.

The model shall be trained using the road dataset of the *Massachusetts Road and Building Detection Datasets* created by (Mnih, 2013) as a part of a Ph.D. thesis. This road dataset contains 1171 3-channel aerial images of the state of Massachusetts at a resolution of 1500 × 1500 pixels where each image covers an area of 2.25 km² (Mnih, 2013). For each aerial image, the dataset contains a target map (ground truth mask) of the roads on that image. The target maps are single-channel Boolean images comprising the classes *road* and *not-road* and were created using OpenStreetMap road data (Mnih, 2013).

At the time of writing, no similar, openly available dataset of roads in Cologne could be found. Therefore, to conduct a change detection in Cologne, a new test dataset must be created. The aerial imagery is obtained as open data provided by the state of Nordrhein-Westfalen in the form of an OGC WMS. This service offers historical digital imagery between the years 1998 and 2020. As no historical road data for the year 1998 could be found, the target masks are created by manual delineation using the aerial imagery for reference. This results in a small number of well-delineated target maps for the years 1998 and 2019.

To train the model the Massachusetts road dataset shall be split up into a training, validation, and test set. As the final use of the model is to segment roads in Cologne, the Massachusetts test set is used to evaluate the performance of the model on very similar data to the data the model has been trained on, before applying the learned parameters to a new setting. Due to the large size of the individual raw images (the 3-channel aerial images in particular), using them as direct input to the network would require a very large amount of computing power and resources. Therefore, as a preprocessing step, the images are either cropped or resized to a more workable resolution (e.g., 256 × 256 pixels or 512 × 512 pixels). Further preprocessing steps are needed to read and transform the images into an appropriate numerical representation for the model. Similarly, the model's output is postprocessed to obtain the predicted target masks in image form. Many factors such as the initialization, learning rate, optimizer, stopping condition, and architectural parameters may influence the performance of the model on the test set (Jakubovitz et al., 2019; Kawaguchi et al., 2017; Lang et al., 2019; Yu & Zhu, 2020). Such hyperparameters shall be evaluated and adjusted during the *Testing, Evaluation, and Validation* phase of the project.

Finally, change detection of the road network in Cologne is conducted by using the trained model to generate target masks for the Cologne test set aerial images. The resulting masks and their accuracy metrics are compared to each other and evaluated to analyze the quality of the automatic change detection. To determine the accuracy of the predicted masks, the dice coefficient or a similar metric can be used (Z. Wang et al., 2020). As the ground truth masks were created manually, the time and effort that the automatic segmentation takes can be directly compared to the manual approach.

## Expected Results

As of the time of writing, the data analysis and finetuning steps have not yet been conducted. Therefore, in this section, the expected results are briefly discussed.

The U-Net model is expected to enable an automated change detection of roads in Cologne, where the changes in the road network are prominent in a visual comparison of the 1998 and 2019 target masks. The target masks predicted by the model are expected clearly outline roads, easily detectable in the aerial images. The model may have some difficulty predicting obscured (e.g., by trees, shadows, or vehicles) sections of the road or roads with varying colors of asphalt or no asphalt. The model is expected to perform better on the Massachusetts road test set than on the Cologne road test set. Although the generalization of learnings to new environments should be possible, the similarity of the test set to the training data will likely be a large advantage for model performance. Furthermore, the final version of the model with adjusted hyperparameters is expected to outperform the earlier iterations on the Cologne test set. Although the general accuracy of the model's segmentation is not expected to be better than the accuracy of manual delineation, the factors of time and effort will most likely be significantly lower using the automated approach (H. Wang et al., 2017).

# References


Gomes, V. C. F., Queiroz, G. R., & Ferreira, K. R. (2020). An overview of platforms for big earth observation data management and analysis. In *Remote Sensing* (Vol. 12, Issue 8). MDPI AG. https://doi.org/10.3390/RS12081253

Jakubovitz, D., Giryes, R., & Rodrigues, M. R. D. (2019). *Generalization Error in Deep Learning*. http://arxiv.org/abs/1808.01174

Kawaguchi, K., Bengio, Y., & Kaelbling, L. (2017). Generalization in Deep Learning. In *Mathematical Aspects of Deep Learning* (pp. 112–148). Cambridge University Press. https://doi.org/10.1017/9781009025096.003

Lang, S., Hay, G. J., Baraldi, A., Tiede, D., & Blaschke, T. (2019). GEOBIA achievements and spatial opportunities in the era of big earth observation data. In *ISPRS International Journal of Geo-Information* (Vol. 8, Issue 11). MDPI AG. https://doi.org/10.3390/ijgi8110474

Matsuyama, T. (1987). Knowledge-Based Aerial Image Understanding Systems and Expert Systems for Image Processing. *IEEE Transactions on Geoscience and Remote Sensing*, *GE-25*(3), 305–316. https://doi.org/10.1109/TGRS.1987.289802

Mnih, V. (2013). *Machine Learning for Aerial Image Labeling*.

Ronneberger, O., Fischer, P., & Brox, T. (2015). *U-Net: Convolutional Networks for Biomedical Image Segmentation*. http://arxiv.org/abs/1505.04597

Sudmanns, M., Tiede, D., Lang, S., Bergstedt, H., Trost, G., Augustin, H., Baraldi, A., & Blaschke, T. (2020). Big Earth data: disruptive changes in Earth observation data management and analysis? In *International Journal of Digital Earth* (Vol. 13, Issue 7, pp. 832–850). Taylor and Francis Ltd. https://doi.org/10.1080/17538947.2019.1585976

Wang, H., Wang, Y., Zhang, Q., Xiang, S., & Pan, C. (2017). Gated convolutional neural network for semantic segmentation in high-resolution images. *Remote Sensing*, *9*(5). https://doi.org/10.3390/rs9050446

Wang, Z., Wang, E., & Zhu, Y. (2020). Image segmentation evaluation: a survey of methods. *Artificial Intelligence Review*, *53*(8), 5637–5674. https://doi.org/10.1007/s10462-020-09830-9

Wu, H., Liu, Q., & Liu, X. (2019). A review on deep learning approaches to image classification and object segmentation. *Computers, Materials and Continua*, *60*(2), 575–597. https://doi.org/10.32604/cmc.2019.03595

Yu, T., & Zhu, H. (2020). *Hyper-Parameter Optimization: A Review of Algorithms and Applications*. http://arxiv.org/abs/2003.05689