Detecting Street Infrastructure Elements from Street-Level Imagery

Graduate

André von Aarburg

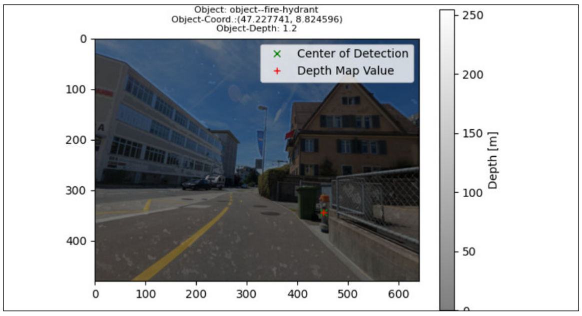
Introduction: OpenStreetMap (OSM) is a global civil society project that collects geospatial data from the ever changing physical and anthropogenetic environment and is used in many domains, like e.g. urban planning. This manual mapping process is time consuming and needs to be repeated in order to keep the database up-to-date. This thesis presents a software application that uses street-level imagery and AI-based computer vision techniques to automatically detect and locate street infrastructure elements by the example of three targets: trash cans, fire hydrants and streetlights (fig. 1).

Approach / Technology: Google Street View paved the way of street-level imagery projects like Mapillary and Panoramax. These two open platforms provide the data basis for this application. While deep learning technologies advanced image recognition, modern algorithms based on Convolutional Neural Networks (CNN) and transformers have improved the ability to detect objects in images with high accuracy, even in cluttered images. Combined with well-known computer vision concepts that allow accurate image reconstruction including depth information (using metadata from the GPS position of the camera) these advances form the basis for the design of such a system. The Python-based application integrates a state-of-the-art instance segmentation model, a Structure from Motion (SfM) algorithm, and spatial clustering. To start the detection of the three targets, an Area of Interest (AOI) has to be defined. A trained YOLO instance segmentation model detects the targets and creates bounding boxes and segmentation masks for each target present in an input image (fig. 1). Then the system processes the images in the AOI to reconstruct them with SfM. With this information, the centroid of the segmentation masks of the targets are used to localize the coordinates (fig. 2). Targets are localized for each detection in an image, so spatial clustering using DBSCAN is required as a final step to remove target ambiguity. Finally, the clustered target coordinates are output to a GeoJSON file for seamless integration and further processing in OSM.

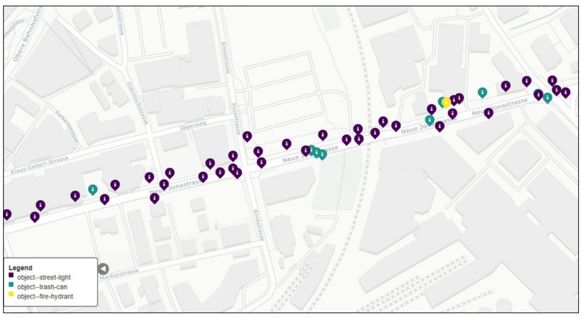
Result: The application showed robust results and allows for further customization of other street infrastructure elements. It has been validated against manually collected OSM data of targets along the Neue Jonastrasse in Rapperswil-Jona (fig. 3). To evaluate the performance of the application, standard metrics such as precision and recall, Sorensen-Dice Coefficient (both defined based on spatial proximity), and the Earth-Mover and Chamfer-Distance metrics were used. Although the system achieved a precision of up to 0.69 for certain targets (streetlights), it suffers from low recall values, which also leads to high values of the implemented distance metrics. This is mainly due to the instance segmentation model, which may misidentify a target that is then localized incorrectly in the downstream SfM pipeline. Additionally, the quality of the depth map provided by the SfM pipeline leaves room for improvement, as regions with low feature intensity, such as pavement, provide noisy depth information, leading to inaccurately localized targets. Despite these drawbacks, the visualization results show high accuracy for correctly predicted targets, providing hints to the "human in the loop". Future development could replace the SfM with monocular depth estimation to improve the localization. Also the runtime performance could be improved this way, which is currently 7 hours for 900 images.

Image along a defined area of interest with detected and segmented targets trash can, fire hydrant and street light.
Own presentment

Aggregated image after localization of the fire hydrant, including depth and coordinate information of the target.
Own presentment



Result of detected and localized streetlights (purple), trashcans (cyan) and hydrants (yellow) in Rapperswil-Jona.
Own presentment



Advisor
Prof. Stefan F. Keller

Co-Examiner
Claude Eisenhut, Eisenhut Informatik, Burgdorf, Bern

Subject Area
Data Science, Computer Science