(Semi)-automatic semantic segmentation ground truth labeling

Degree programme: Master of Science in Engineering
Thesis advisor: Prof. Marcus Hudritsch
Expert: Dr. Harald Studer (OPTIMO MEDICAL AG)

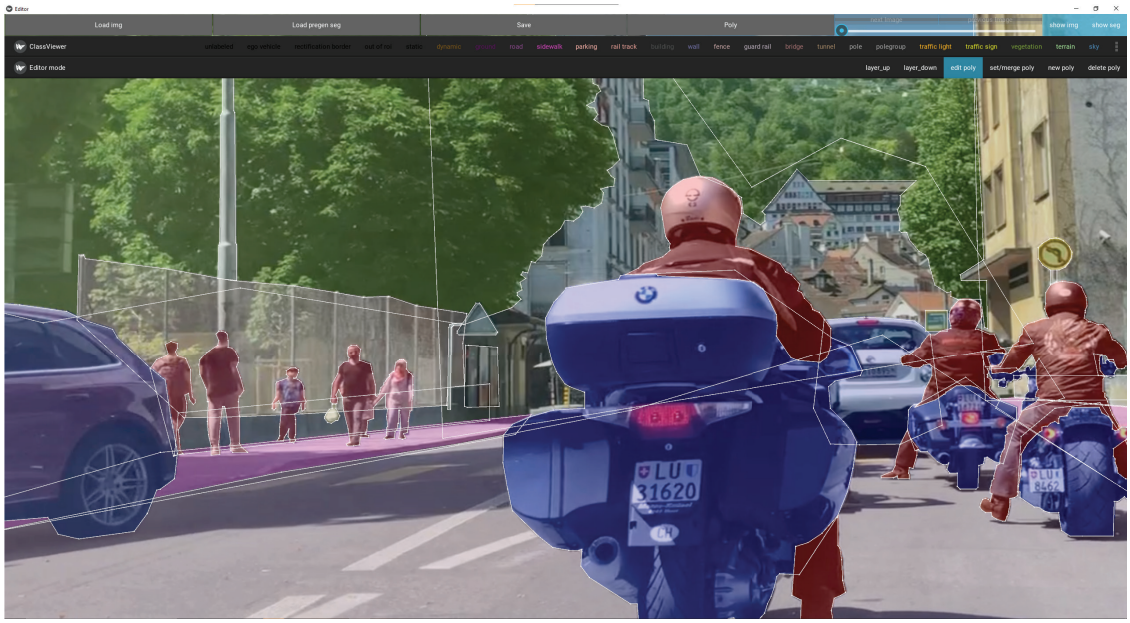
The goal of this thesis is, given a neural network architecture from the previous project, to create a tool for labeling further data for semantic segmentation of street scenes. To analyze the quality of the data generated and the benefit, if any, it has on the overall performance.

Semantic segmentation is the process of assigning each pixel of an input image a class to which it belongs. In the domain of the thesis a few of the more important such classes are, ‚car', ‚person', ‚road', ‚traffic sign' and so on. After this labeling one has a so called segmentation mask, which could be used as input for a driving AI. The well known company Tesla for instance does, among other things, use semantic segmentation for its autonomous vehicles. The problem with semantic segmentation is that state of the art results can only be achieved by neural networks, however the generation of so called ground truth data which is required for the training of the network is a painstaking process involving hours and hours of labeling by hand.

The goal of this thesis is to address the issue of labeling, to this end different possibilities for assisted ground truth labeling will be explored. The core idea is that given an already trained network we should be able to use it to assist in the labeling of further data. This data will of course be of a lower quality than fully hand labeled data, hence one of the core questions is: Is the data good enough to increase the networks performance?
One might wonder why this trade-off might be a desirable thing, the reason is simple, labeling images fully by hand even with (somewhat) sophisticated tools takes a long time, 25-70 minutes. Now with the quantity of data neural networks require for proper training this equates to weeks and weeks of labeling data for even only a moderately sized dataset. If this process could be sped up by several times without too much of a hit to the quality, that would likely result in more and better networks.

To achieve all of this a network trained and developed in a previous project was integrated into a tool chain for training, inferencing and labeling.

Thomas Michael Voegelin
079 397 54 98



Detailed labeling using the developed tool