Improving ground truth for CNNs through Uncertainty and a Human-in-the-loop

Motivation

One of the most important things in the area of machine learning is to have trustworthy ground truth to train the network. For image labeling and other tasks often humans are generating this ground truth. But with increasing amounts of data it is getting extremely time consuming to perform this work. Especially when there are up to several billions of voxels, it is unfeasible to label every single voxel manually. To reduce this effort often a threshold-based segmentation algorithm is used to create the ground truth. Using a segmentation algorithm to create the ground truth adds a lot of uncertainty and will bias the whole training of a CNN.

Goal

The task is to implement an algorithm which is based on multiple segmentation methods parameterized from a human. The human will after every prediction of the network give feedback for a small number of subvolumes and improve by a manual controlled segmentation algorithm the ground truth for these subvolumes. The subvolumes for the feedback should be selected according to an uncertainty metric (Prediction result, statistics, …). This procedure will be repeated several times (Human in the loop) until the user is satisfied by the prediction results of the CNN.

Kontakt

Patrick Weinberger (patrick.weinberger@fh-wels.at)