Semantic Instance Segmentation for Autonomous Driving

Wenjie Niu

May 23, 2018

This is a summary of Semantic Instance Segmentation for Autonomous Driving, an article in *CVPR*,2017[1].

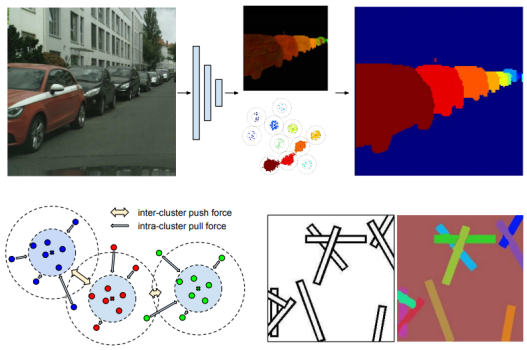

Figure 1: Top: the network maps each pixel into feature space, in which each object can be easily clustered with a fast clustering approach. Bottom left: intra-cluster pull and inter-cluster push forces. Bottom right: our method can handle complex occlusions, useful for pick and place tasks.

1 Occlusion handling

A key strength of our method is its ability to handle complex occlusions. Detect-and-segment approaches require an objects segmentation mask to be unambiguously extracted from its bounding box. This assumption is problematic for certain tasks. Consider a pick-and-place task where overlapping stick-like objects need to be segmented as in fig1. When two sticks overlap like two crossed swords, their bounding boxes are highly overlapping. Given only a detection in the form of a bounding box, it is exceedingly hard to unambiguously extract a segmentation mask of the indicated object. In contrast to methods that rely on bounding boxes, our method treats the image holistically and can learn to reason about occlusions.

2 Scene understanding for autonomous driving

We test our loss function on the challenging Cityscapes dataset, a multi-class semantic instance segmentation benchmark. To cope with the multi-class problem, we apply our loss function independently on each semantic class so that instances from different classes are free to occupy the same feature space. The semantic segmentation masks are obtained with the ResNet-38 network. The same architecture, pretrained on Cityscapes semantic segmentation, is also adopted for our instance segmentation network. We train the model on the 2975 training images, resized to 768x384 and use Adam with learning rate of 1e-4 on a NVIDIA Titan X. With our loss we achieve competitive results on the Cityscapes leaderboard, see table 1.

	AP	AP0.5	AP100m	AP50m
InstanceCut	13.0	27.9	22.1	26.1
DWT	15.6	30.0	26.2	31.8
Shape-aware	17.4	36.7	29.3	34.0
Pixelwise DIN	20.0	38.8	32.6	37.6
Ours	17.5	35.9	27.9	31.0

Table 1: Segmentation results of best performing entries on the test set of the Cityscapes instance segmentation benchmark.

References

[1] Bert De Brabandere, Davy Neven, and Luc Van Gool. Semantic instance segmentation for autonomous driving. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, July 2017.