KI DATA TOOLING
The Data Kit for Automotive AI

Amodal Cityscapes: A New Dataset, its Generation, and an Amodal Semantic Segmentation Challenge Baseline

Jasmin Breitenstein, Tim Fingscheidt,
Technische Universität Braunschweig

Amodal Perception

Humans have the ability to anticipate the full shape of (partially) occluded objects. Typical perception methods for automated driving lack this ability, leading to worse detection performance in the presence of occlusions, thus resulting in corner cases. Moreover, other corner case types can form behind occlusions.

Figure 1: Example image (input) and corresponding annotations (output) from „Amodal Cityscapes"

Amodal Cityscapes Dataset

We create the Amodal Cityscapes dataset, which is based on the standard Cityscapes dataset. With the new dataset, we can train an amodal semantic segmentation with amodal ground truth without the expensive labeling of real data.

	training	val	test
#images	2900	75	500
#masks	2900	75	500
#target images	2900	75	500
#source images	2900-1	75-1	500-1
#pasting occluders	36303-N	832-N	6438-N

Table 1: Split of the Amodal Cityscapes dataset

Table 1 shows the split of our dataset. We paste occluders according to their original vertical coordinate and sample the horizontal coordinate randomly from the image width. This keeps the context of images intact. Per target image, an occlusion ratio is randomly sampled from [0,0.1], and occluders are inserted until the occlusion ratio is exceeded.

Challenge Baseline Method

To show the merit of our dataset for amodal semantic segmentation, we adapt the amodal semantic segmentation of Purkait et al. [1] to the ERFNet. It is based on a groupwise encoding of the semantic classes into static, traffic, person-like, vehicle-like. Figure 2 shows an example encoding for one pixel. Then, the final softmax layer is replaced by a groupwise softmax layer that predicts the visible group and per group the most likely semantic class independent of their visibility.

Experimental Results

Results of the standard ERFNet and the amodal semantic segmentation (amERFNet) are shown in Table 2, $\mathrm{mIoU}_{\mathrm{inv}}$ measures mIoU on the amodal image parts.

Method	mIoU	$\mathrm{mIoU}_{\mathrm{inv}}$
ERFNet	62.99%	5.00%
amERFNet	62.76%	23.60%

Table 2: Performance of ERFNet and amERFNet on the Amodal Cityscapes test dataset

Figure 3 shows qualitative results of the amERFNet. On both datasets, we see a reasonable hallucination behind occlusions.

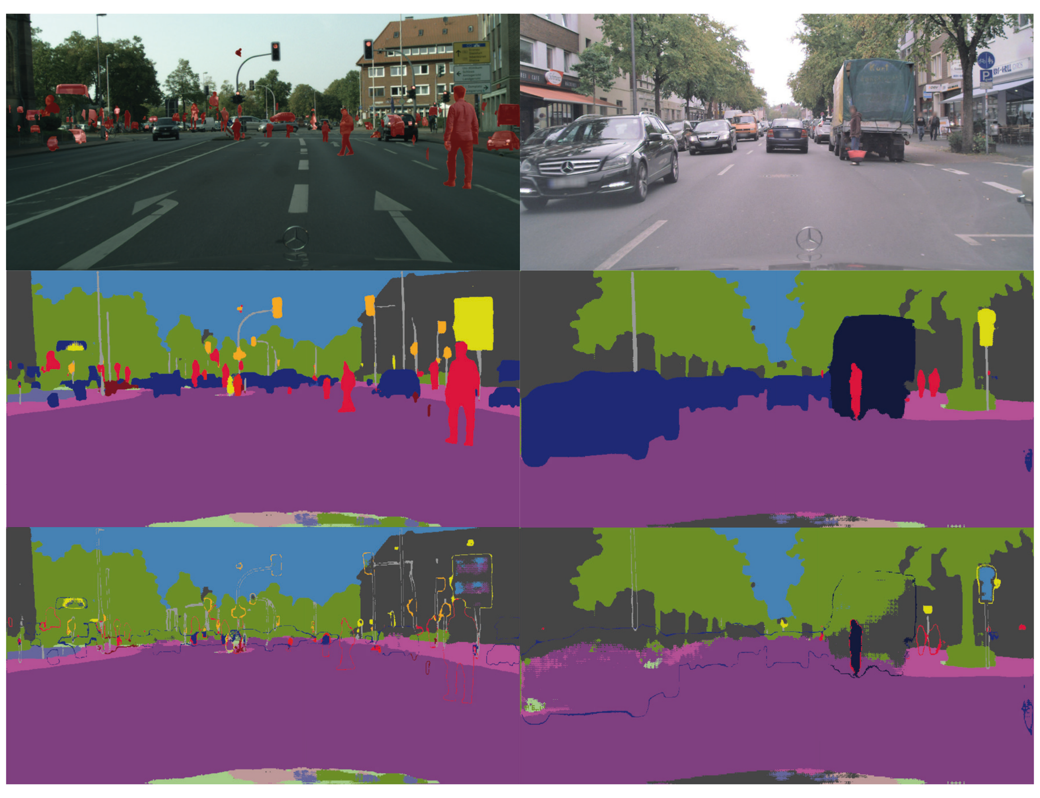

Figure 3: Example image and corresponding predictions from the Amodal Cityscapes test set (left) and the Cityscapes validation set (right)

References:

[1] P. Purkait, C. Zach, and I. D. Reid, „Seeing Behind Things: Extending Semantic Segmentation to Occluded Regions," IROS 2019

[2] J. Breitenstein, T. Fingscheidt: „Amodal Cityscapes: A New Dataset, its Generation, and an Amodal Semantic Segmentation Challenge Baseline," IV 2022

$$\bar{\boldsymbol{y}}_i = (\underbrace{\boxed{0,0,1,0,}}_{\text{Visible group: person-like}}\ \underbrace{\boxed{0,0,0,0,0,1,0,0,0,}}_{\text{Static group: class terrain is present}}\ \underbrace{\boxed{0,0,0,1,}}_{\text{Traffic group: absent}}\ \underbrace{\boxed{0,1,0,}}_{\text{Person-like group: class rider is present}}\ \underbrace{\boxed{0,0,0,0,0,0,1}}_{\text{Vehicle-like group: absent}})$$

Figure 2: Example groupwise encoding for a pixel with visible class rider and invisible class terrain

Technische Universität Braunschweig

For more information contact:
j.breitenstein@tu-bs.de

KI Data Tooling is a project of the KI Familie. It was initiated and developed by the VDA Leitinitiative autonomous and connected driving and is funded by the Federal Ministry for Economic Affairs and Climate Action.

Supported by:
Federal Ministry for Economic Affairs and Climate Action
on the basis of a decision by the German Bundestag

KI FAMILIE

https://www.ki-absicherung-projekt.de   @KI_Familie   KI Familie