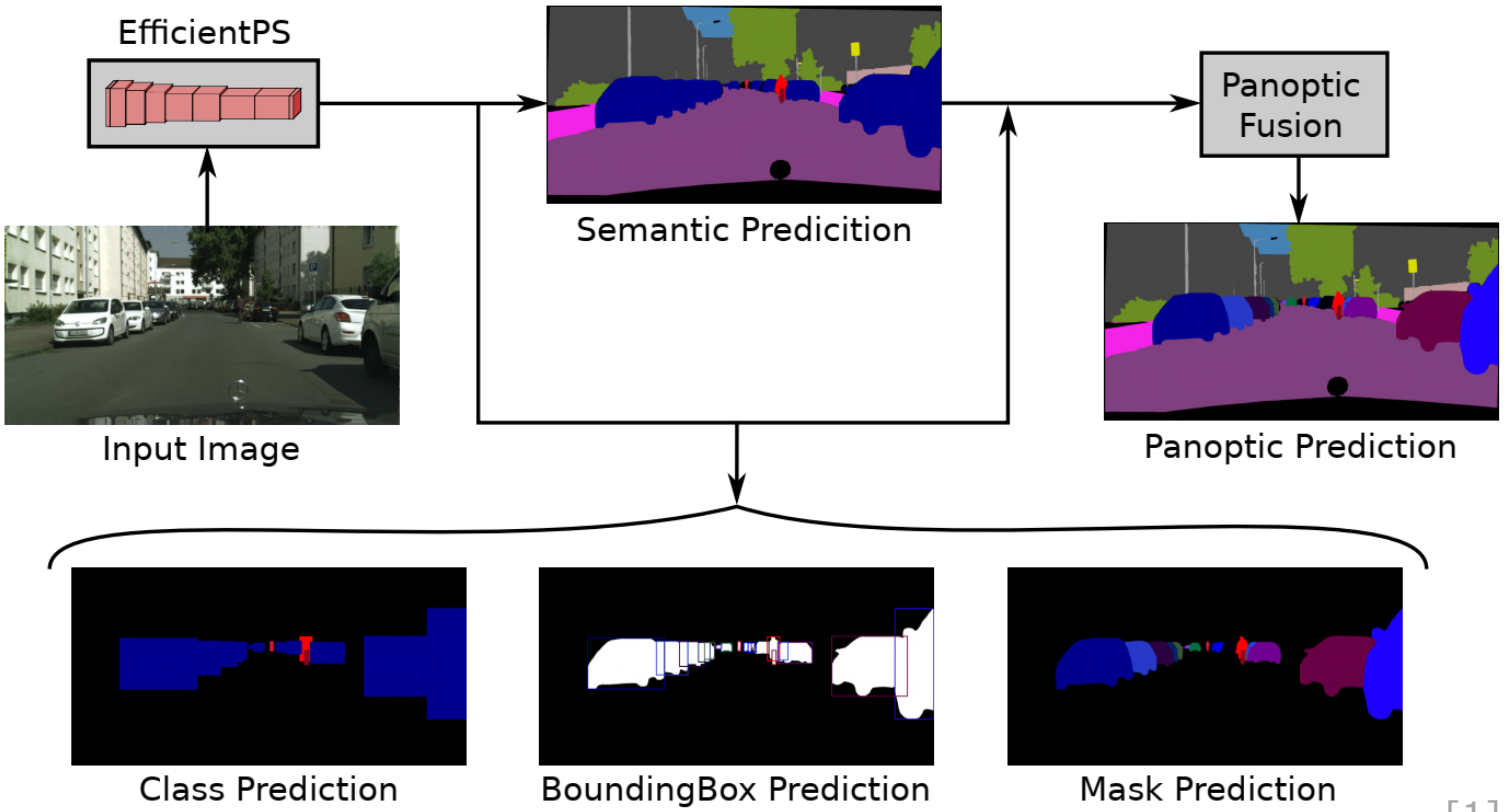
Crop-Aware Panoptic Segmentation

Mehdi Naouar, Ben Bausch, Yannick Vogt

Panoptic Segmentation



[1]

Bounding Box Prediction

- Region Proposal Network proposes anchor boxes
- Fully Connect Layer regresses transformations applied to the anchor to fit the ground truth Bounding Box
- Smooth-L1 Loss

$$L_{\mathrm{BB}}(\Delta_{\mathrm{P}};\Delta_{\mathrm{G}}) = \|\ell_\beta(\boldsymbol{\delta}_{\mathrm{P}}-\boldsymbol{\delta}_{\mathrm{G}})+\ell_\beta(\log\boldsymbol{\omega}_{\mathrm{P}}-\log\boldsymbol{\omega}_{\mathrm{G}})\|_1$$

$$\ell_\beta(z) = \begin{cases} \frac{1}{2\beta}z^2 & |z| \le \beta \\ |z| - \frac{\beta}{2} & \text{otherwise,} \end{cases}$$

Bounding Box Example

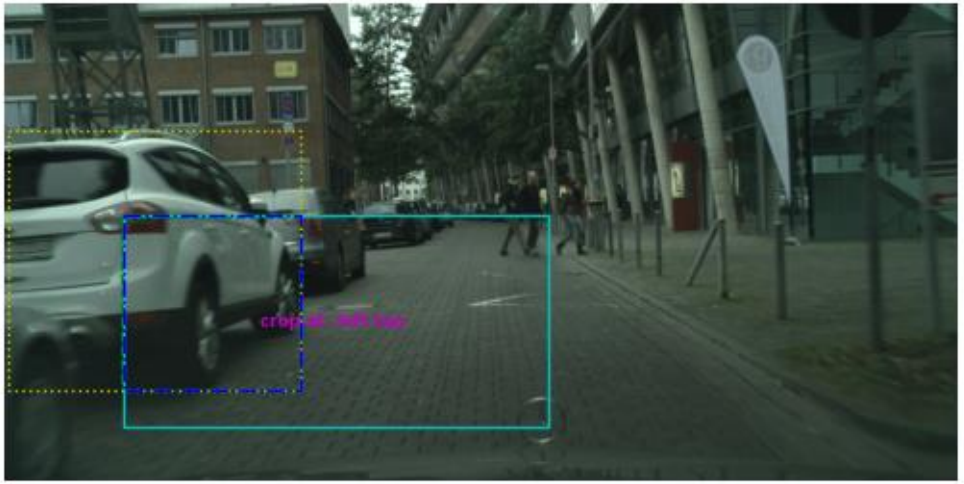
Ground Truth, Anchor Box

Motivation

- Full-size images consume too much memory
- Resized images can’t represent fine structures
- Idea: Train on crops

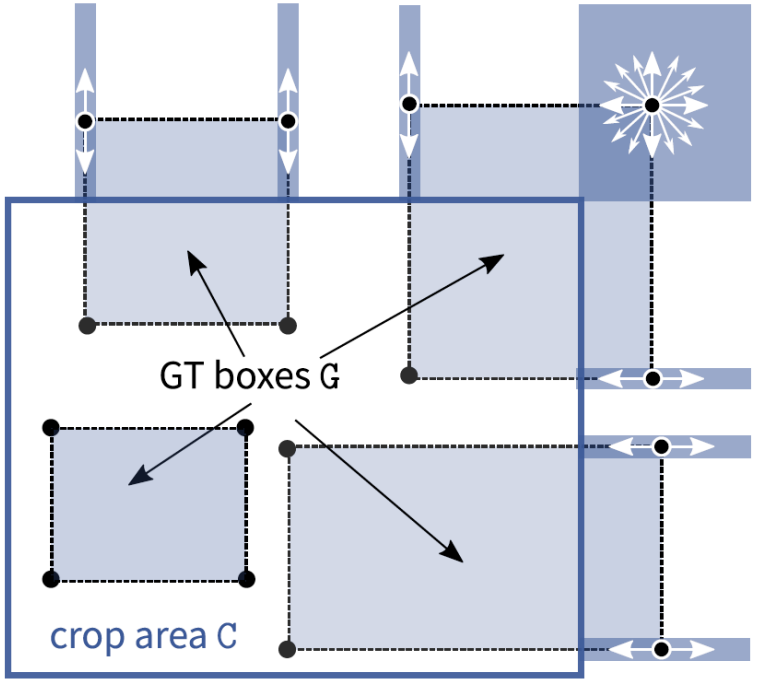
Crop Based Training

- Allows training of larger networks or higher batch size
- Fine structures can be preserved
- Large objects might be truncated



Crop-Aware Bounding Boxes (CABB)

- Network is allowed to "hallucinate"
- Can improve performance on objects which were cropped in training



[2]

Visualization of CABB

Without CABB

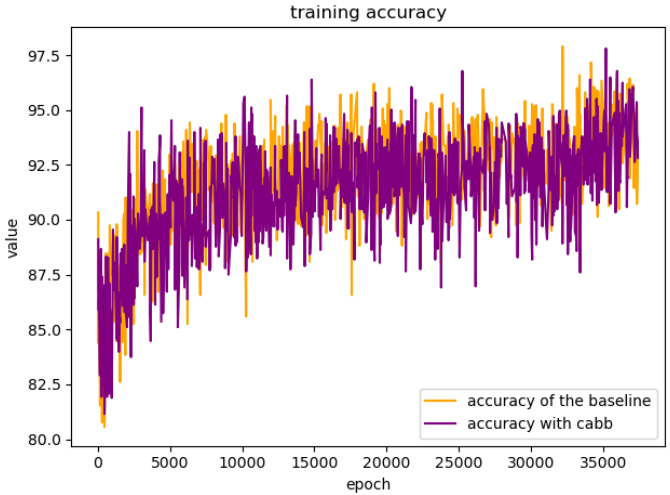
With CABB

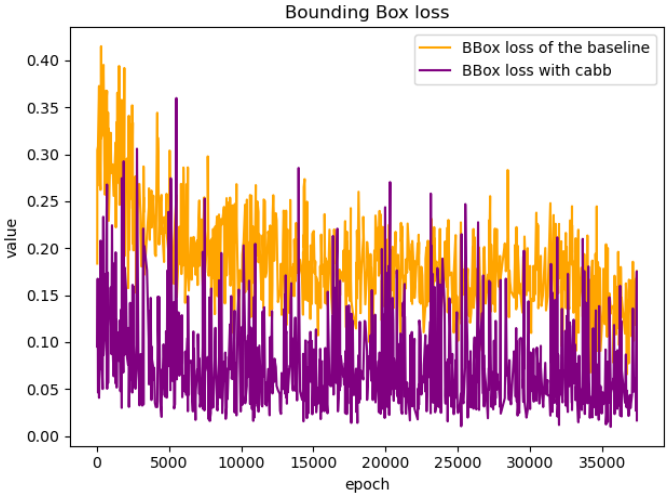
Ground Truth, Prediction, Crop Area, Anchor Box, CABB Target

Training Setting

- 1024*512 Image size, 512*256 Crop size
- 8 Image per Batch on a single GPU
- 100 epochs of training with 2975 training images

Training Results





- 39.5h for baseline vs 20.75h for the baseline

Test Results

- Test models at epoch 100 on 500 full-sized images

Base	PQ	SQ	RQ	N
All	47.6	75.5	60.9	19
Things	42.0	75.1	55.8	8
Stuff	51.7	75.8	64.5	11

CABB	PQ	SQ	RQ	N
All	46.1	74.9	59.2	19
Things	38.2	72.7	52.1	8
Stuff	51.9	76.5	64.4	11

Sources

- [1] Rohit Mohan, Abhinav Valada**. EfficientPS: Efficient Panoptic Segmentation.** arXiv:2004.02307v3, 2021
- [2] Lorenzo Porzi, Samuel Rota Bulò, Peter Kontschieder. **Improving Panoptic Segmentation at All Scales.** arXiv:2012.07717v2, 2021

Thank you for your attention!