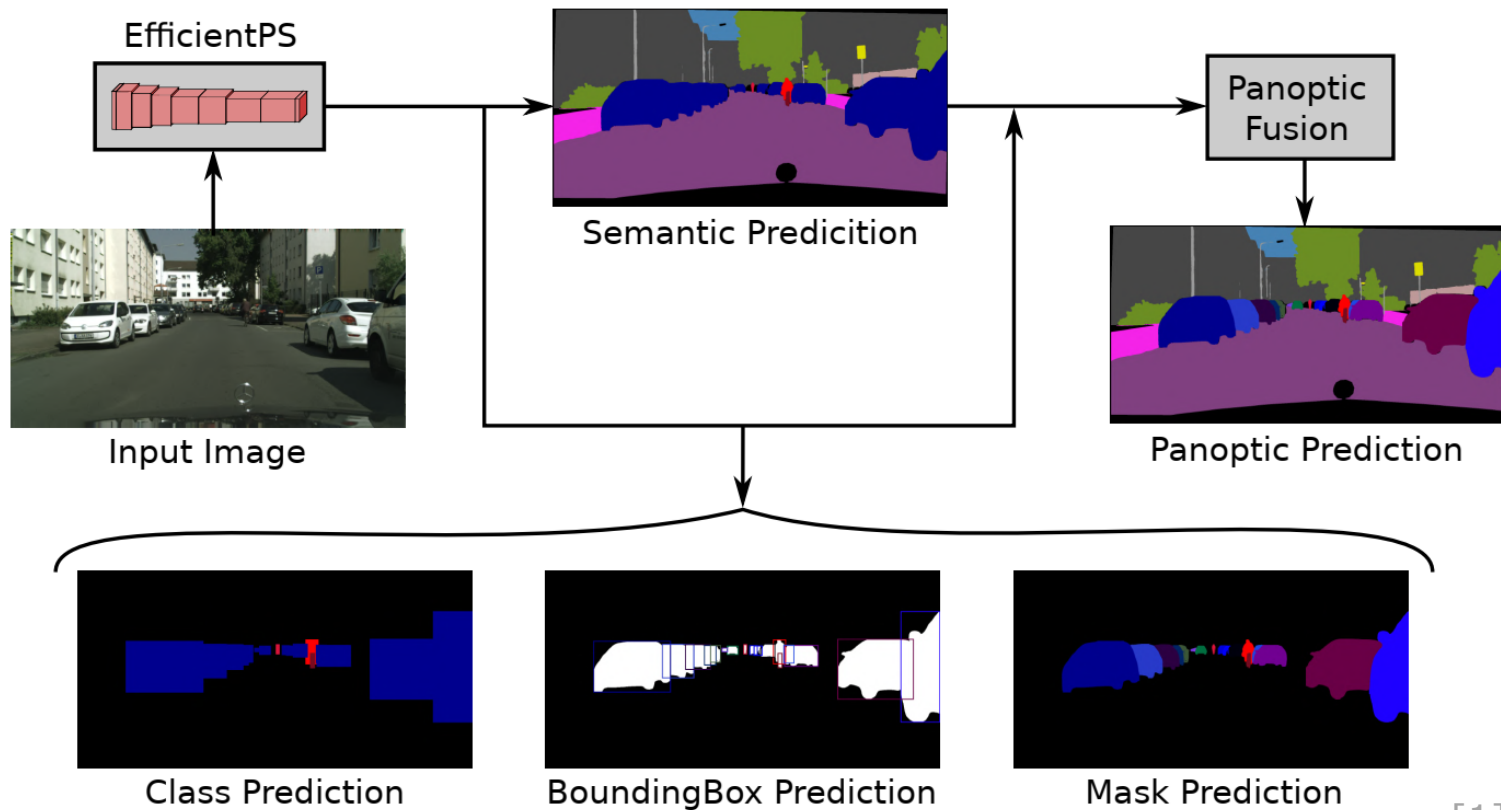


Crop-Aware Panoptic Segmentation

Mehdi Naouar, Ben Bausch, Yannick Vogt



Panoptic Segmentation



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

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

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

Bounding Box Example



Ground Truth, **Anchor Box**

Motivation

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

Crop Based Training

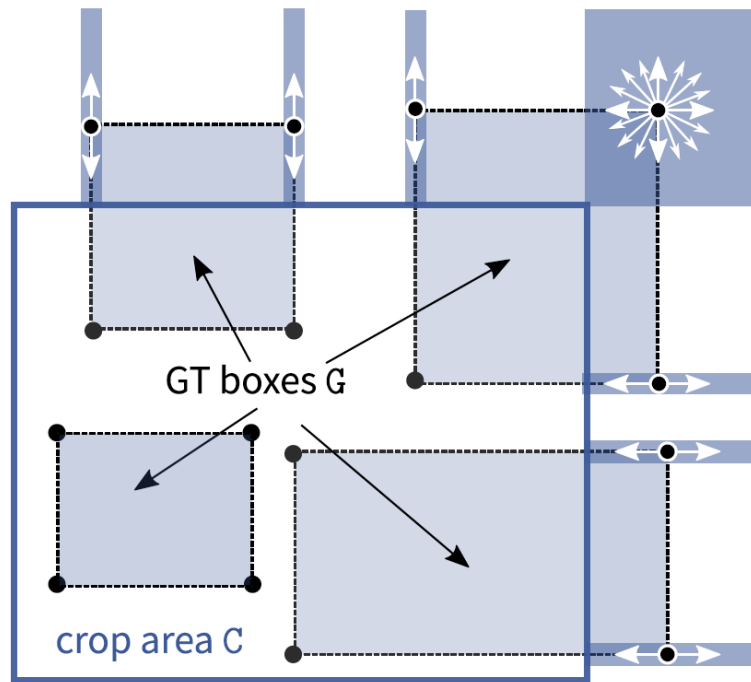
- Allows training of larger networks or higher batch size
- Fine structures can be preserved
- Large objects might be truncated



Cropped Ground Truth,
Original Ground Truth,
Crop Area

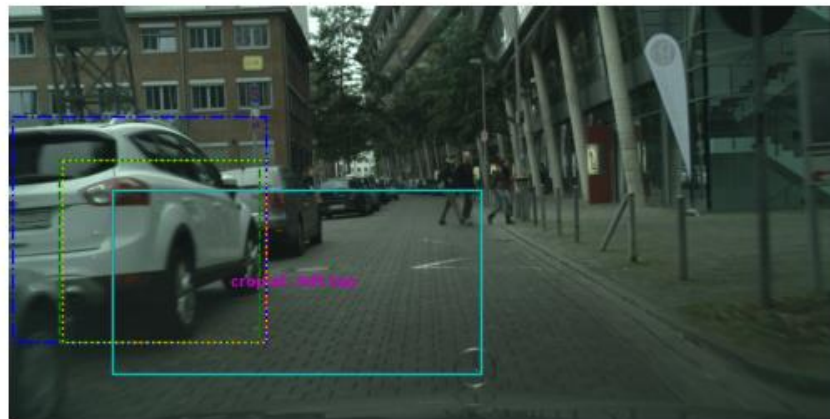
Crop-Aware Bounding Boxes (CABB)

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



[2]

Visualization of CABB



Error without CABB

Error with CABB

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

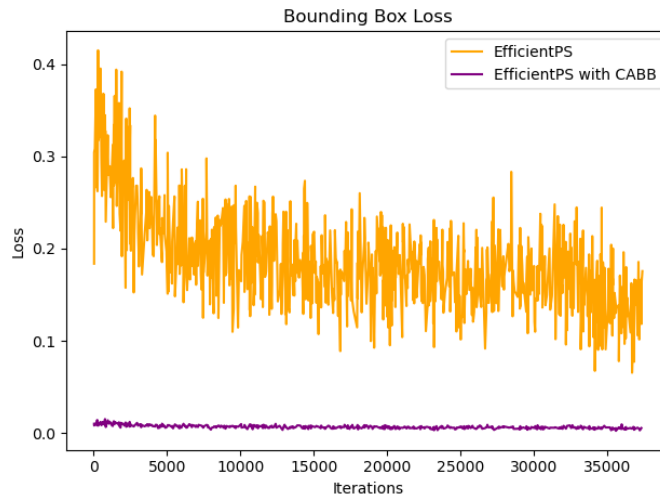
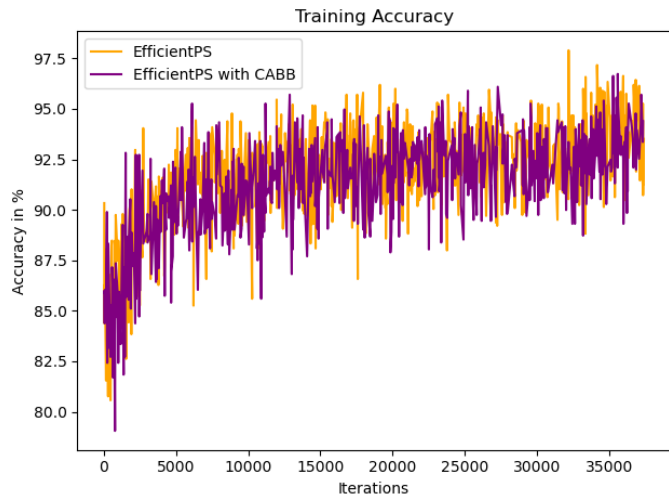
Dataset

- Cityscapes Dataset
 - 2975 training samples
 - 500 validation samples
 - 1525 test samples without ground truth
 - 2048 x 1024 pixels

Training Setting

- 1024 x 512 image size, 512 x 256 crop size
- 100 epochs of training with 2975 training images
- 8 image per batch on a single GPU
- ~37.200 iterations

Training Results



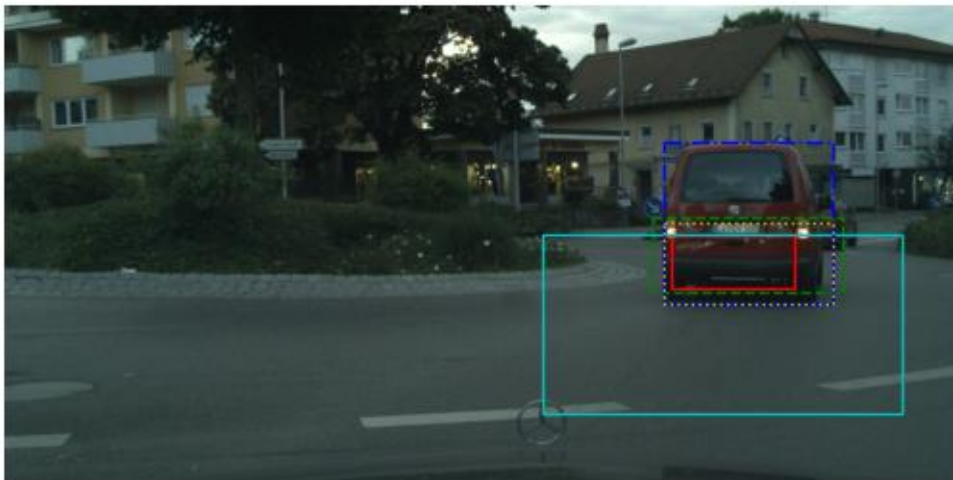
- 39.5h with CABB vs 20.75h without CABB

Validation Results

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

Without CABB	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	35.0	71.5	46.2	19
Things	32.0	72.2	44.2	8
Stuff	37.1	71.0	47.7	11

CABB prediction after Training



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

Panoptic Segmentation



EfficientPS

EfficientPS with CABB

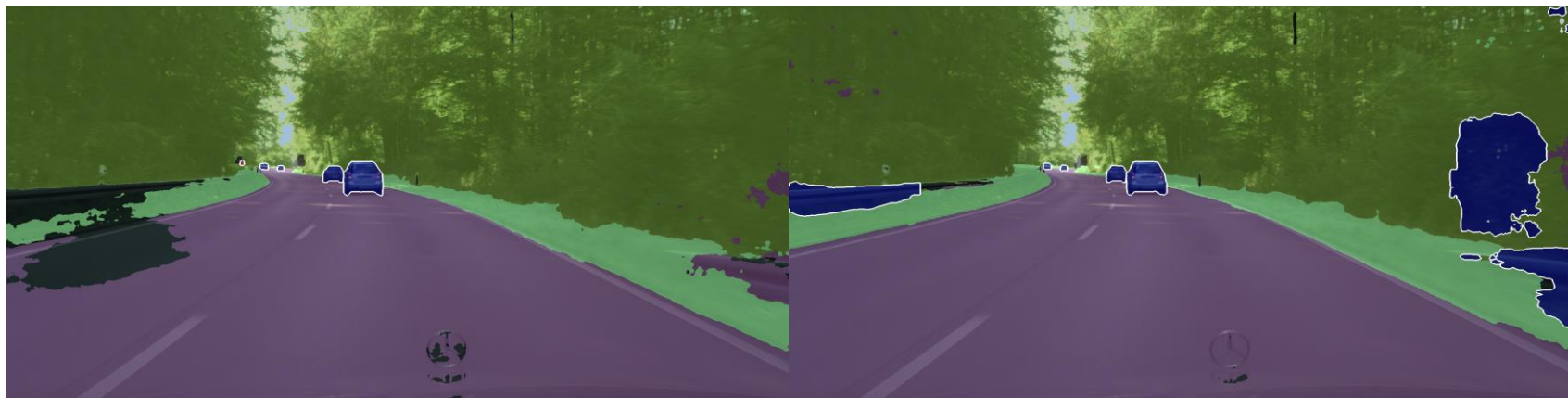
Conclusion

- Massive decrease in performance
- Potential issues:
 - Mistakes in implementation
 - Loss too low in comparison to other losses
 - Network crops objects based on Bounding Boxes

References

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

Thank you for your attention!



EfficientPS

EfficientPS with CABB



EfficientPS

EfficientPS with CABB