HRDA: Context-Aware High-Resolution Domain-Adaptive Semantic Segmentation

Lukas Hoyer, Dengxin Dai, Luc Van Gool • github.com/lhoyer/HRDA

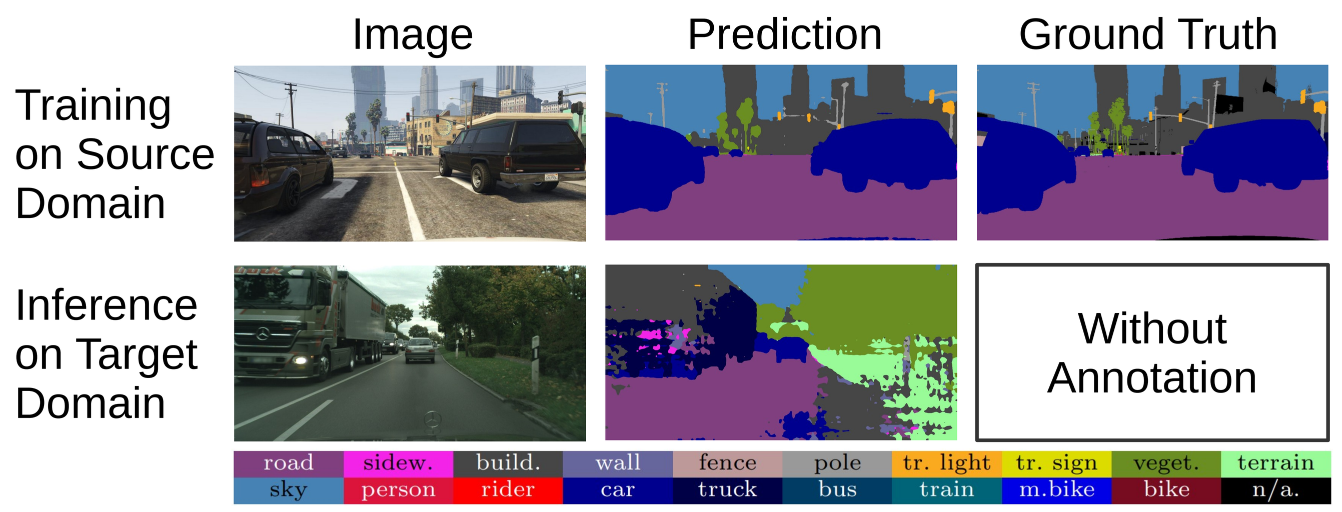
max planck institut informatik

ETH zürich

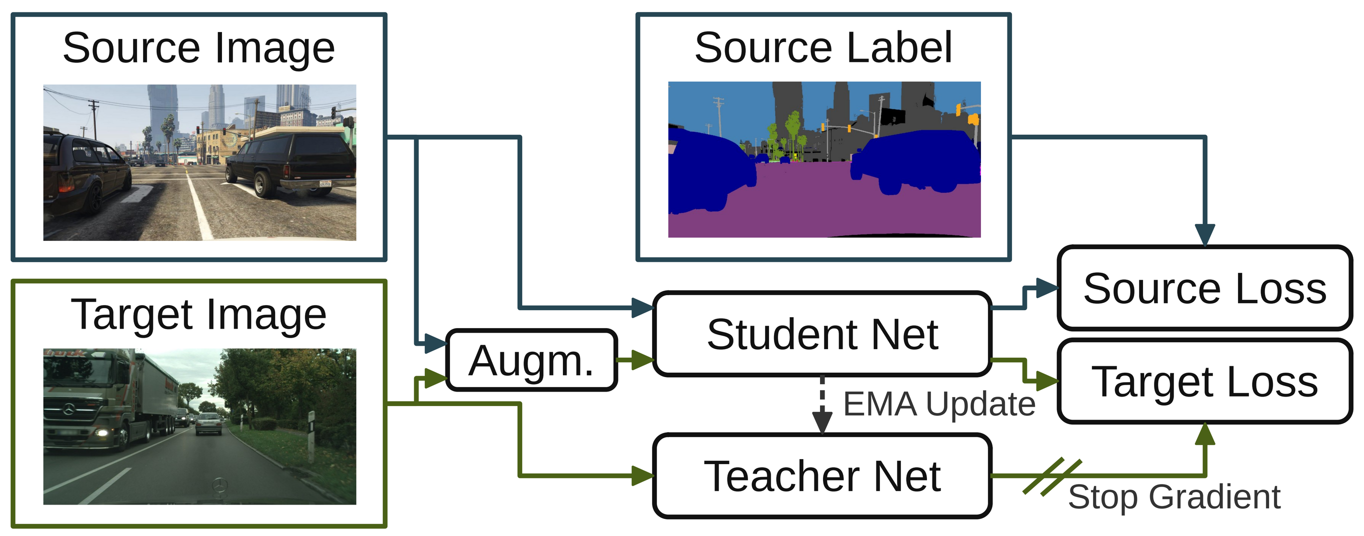
ECCV TEL AVIV 2022

1 Domain Adaptation (UDA)

Problem: Networks trained on one domain often experience a performance drop on another domain



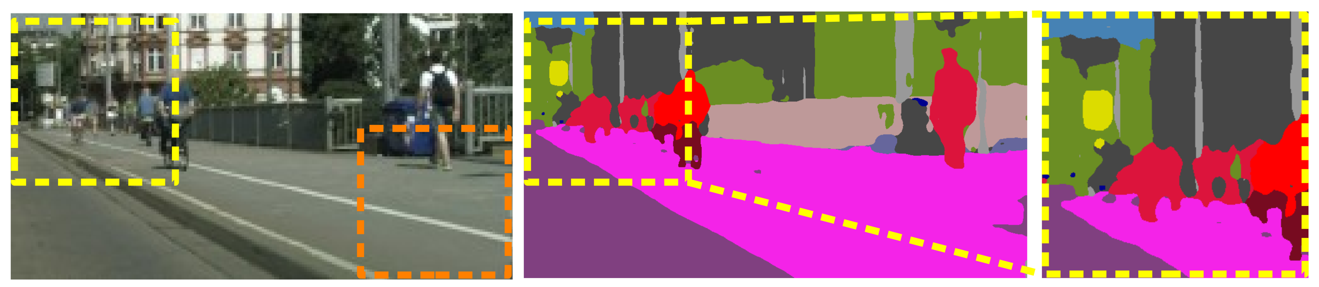
Solution: Adapt network to unlabeled target images, for example using self-training [2,3]



2 HRDA Motivation

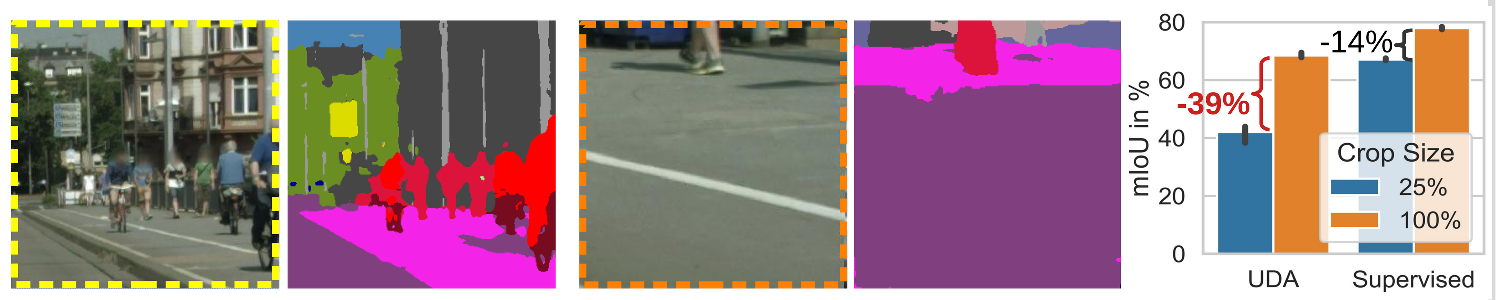
Problem: UDA is more GPU memory demanding than supervised learning

Approach 1: Training with downscaled images



➔ Fails to preserve fine details

Approach 2: Training with random crops



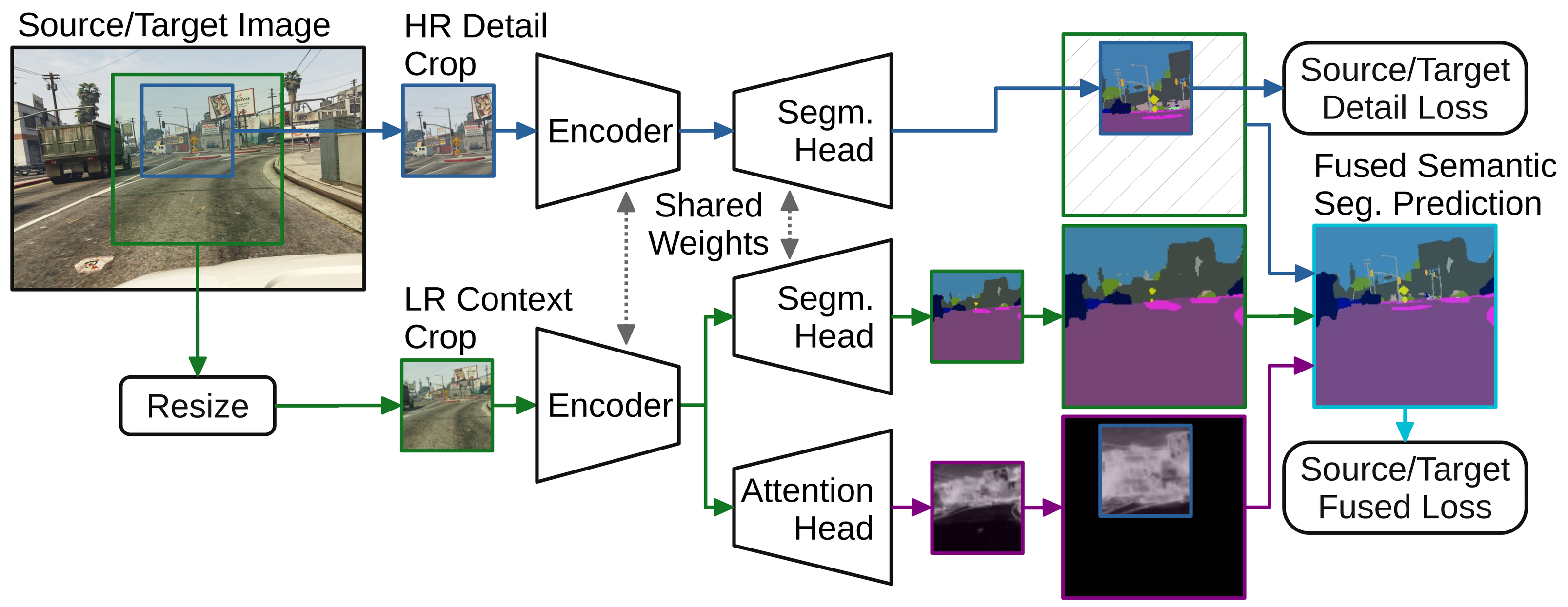
➔ Preserves fine details

➔ Fails to capture context (crucial for UDA)

3 HRDA Method

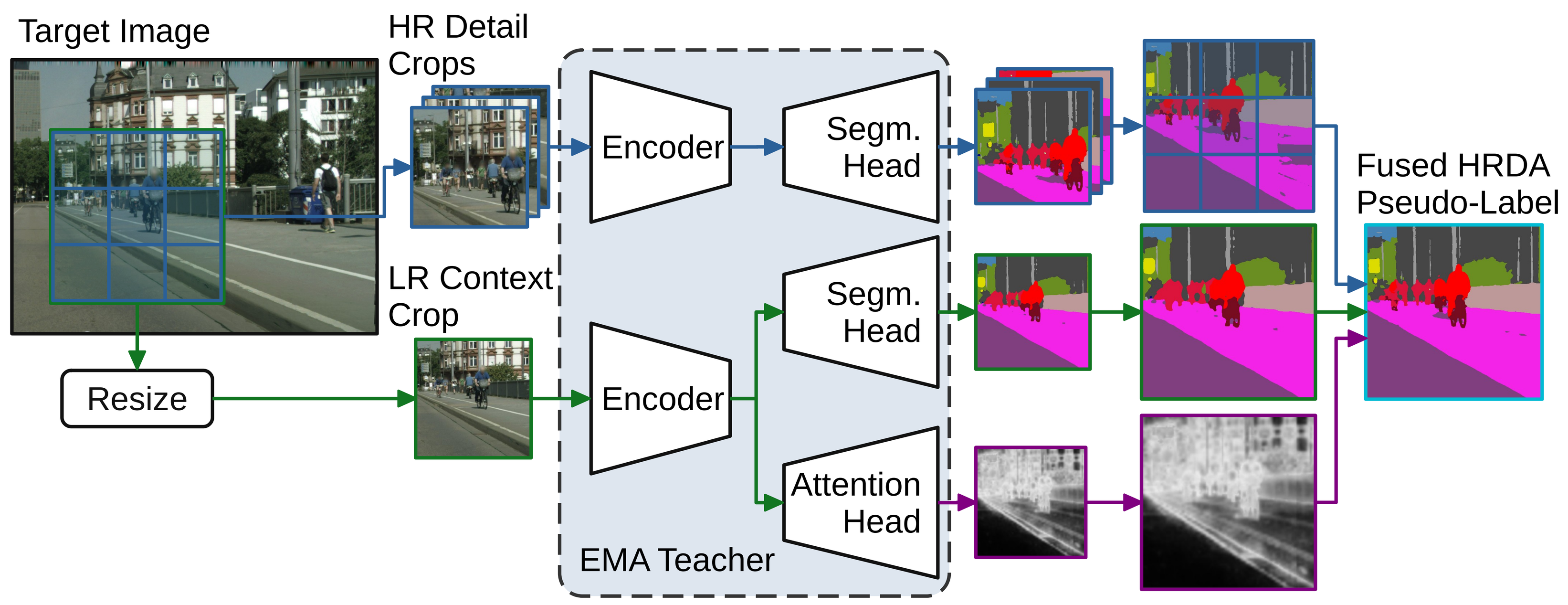
Multi-Resolution Training

- Large low-resolution (LR) crops for long-range context
- Small high-resolution (HR) crops for fine details
- Learned scale attention for scale-dependent adaptation
- Manageable GPU memory consumption

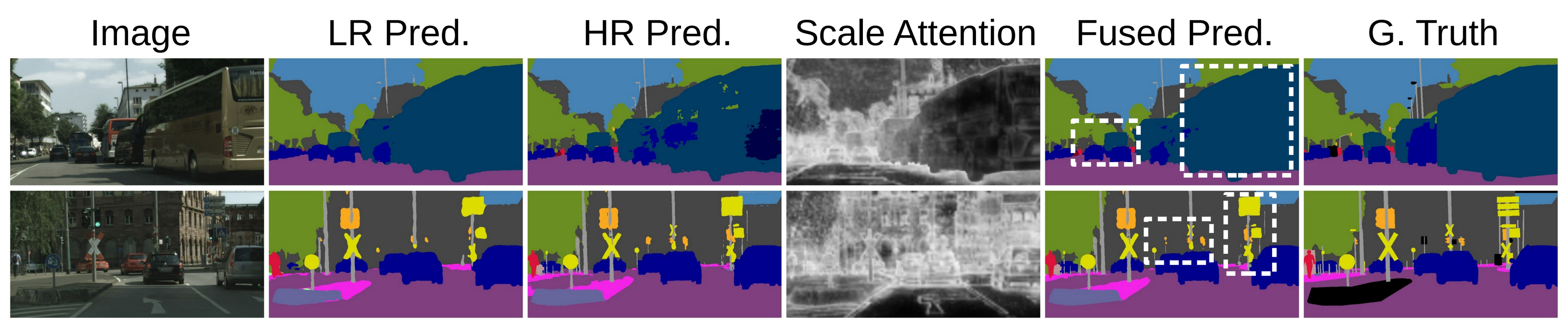


Multi-View Pseudo-Label Generation

- Multi-resolution fusion with scale attention
- Detail context fusion with overlapping sliding window



Multi-Resolution Examples



References

[1] Chen et al. "Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs." PAMI. 2017

[2] Hoyer et al. "DAFormer: Improving network architectures and training strategies for domain-adaptive semantic segmentation." CVPR. 2022.

[3] Tranheden et al. "DACS: Domain Adaptation via Cross-domain Mixed Sampling." WACV. 2021.

[4] Tsai et al. "Learning to adapt structured output space for semantic segmentation.". CVPR. 2018.

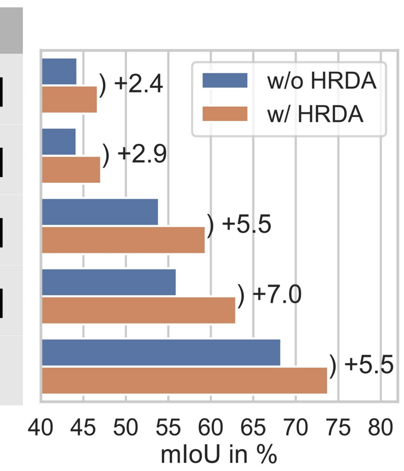
[5] Vu et al. "Advent: Adversarial entropy minimization for domain adaptation in semantic segmentation." CVPR. 2019.

[6] Zhang et al. "Prototypical pseudo label denoising and target structure learning for domain adaptive semantic segmentation." CVPR. 2021

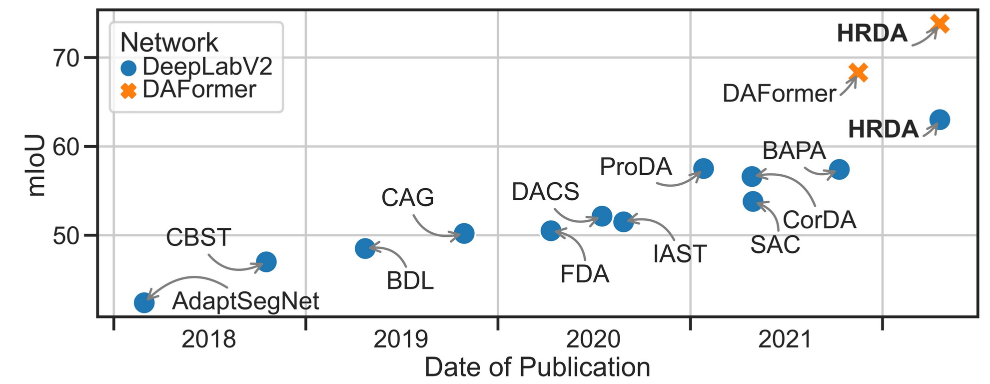
5 Evaluation on GTA➔Cityscapes

Comparison with State-of-the-Art

UDA Method	Network
Entropy Min. [5]	DeepLabV2 [1]
Adversarial [4]	DeepLabV2 [1]
DACS [3]	DeepLabV2 [1]
DAFormer [2]	DeepLabV2 [1]
DAFormer [2]	DAFormer [2]



➔ HRDA improves the performance across different UDA methods and networks



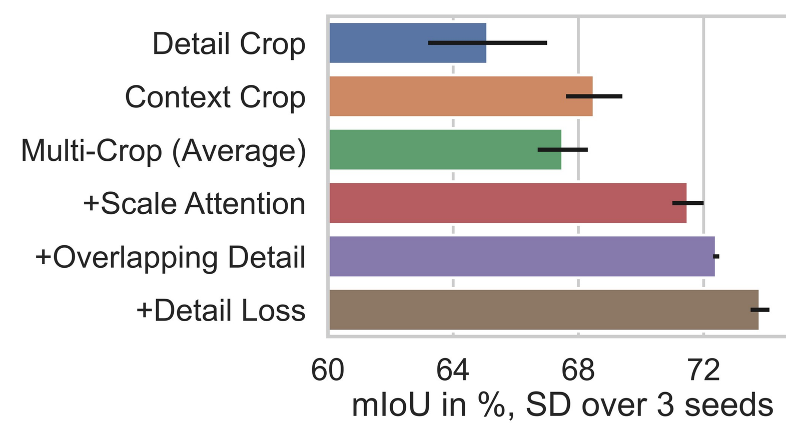
➔ HRDA improves the State-of-the-Art by 5.5 mIoU

	Road	S.walk	Build.	Wall	Fence	Pole	T.Light	T.Sign	Veget.	Terrain	Sky	Person	Rider	Car	Truck	Bus	Train	M.bike	Bike
DAFormer	96	71	89	54	49	49	56	60	90	50	92	72	46	92	69	80	67	57	62
HRDA	96	74	91	62	51	57	64	69	91	48	94	79	53	94	84	86	76	64	68

Class-Wise IoU in %

➔ HRDA improves small objects and difficult classes

Ablation Study



➔ HRDA improves over training with only context/detail crop

➔ Scale attention is crucial

➔ Overlapping detail pseudo-labels increase robustness

➔ Additional detail loss helps learning HR features

UDA Method	GPU Memory	mIoU in %
Naive HR (75% crop size)	22.0 GB	70.0 ± 1.2
HRDA (75% crop size)	13.5 GB	71.3 ± 0.3
HRDA (Full crop size)	22.5 GB	73.8 ± 0.3

➔ HRDA outperforms naive HR training

Example Predictions