Curriculum Domain Adaptation for Semantic Segmentation of Urban Scenes

Yang Zhang, Philip David**, Boqing Gong*

*Center for Research in Computer Vision, University of Central Florida, Orlando, FL 32816
**Computational and Information Sciences Directorate, U.S. Army Research Laboratory, Adelphi, MD 20783

yangzhang@knights.ucf.edu, philip.j.david4.civ@mail.mil, bgong@crcv.ucf.edu

UCF CENTER FOR RESEARCH IN COMPUTER VISION

ARL

ICCVW17
International Conference on Computer Vision 2017

Motivation

Segmentation model trained on synthetic data does not work well on real images

Laborious to annotate real training images
Synthetic training images mismatch with real images

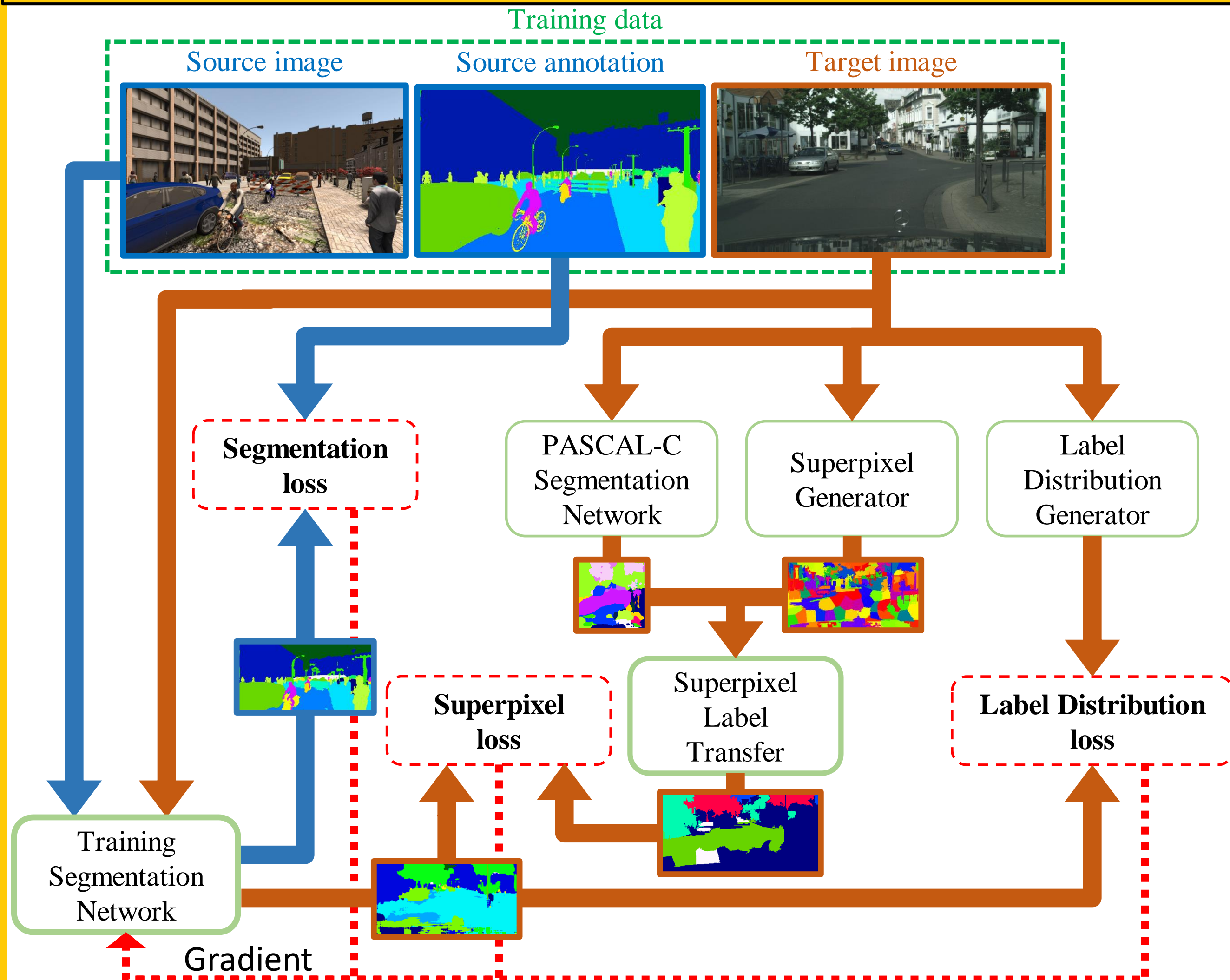

Visual difference caps model performance

Goal: Domain adaptation for semantic segmentation:

Boost performance of models trained on synthetic images

Approach

Training data

Source image   Source annotation   Target image

Segmentation loss

PASCAL-C Segmentation Network

Superpixel Generator

Label Distribution Generator

Superpixel loss

Superpixel Label Transfer

Label Distribution loss

Training Segmentation Network

Gradient

Curriculum Domain Adaptation

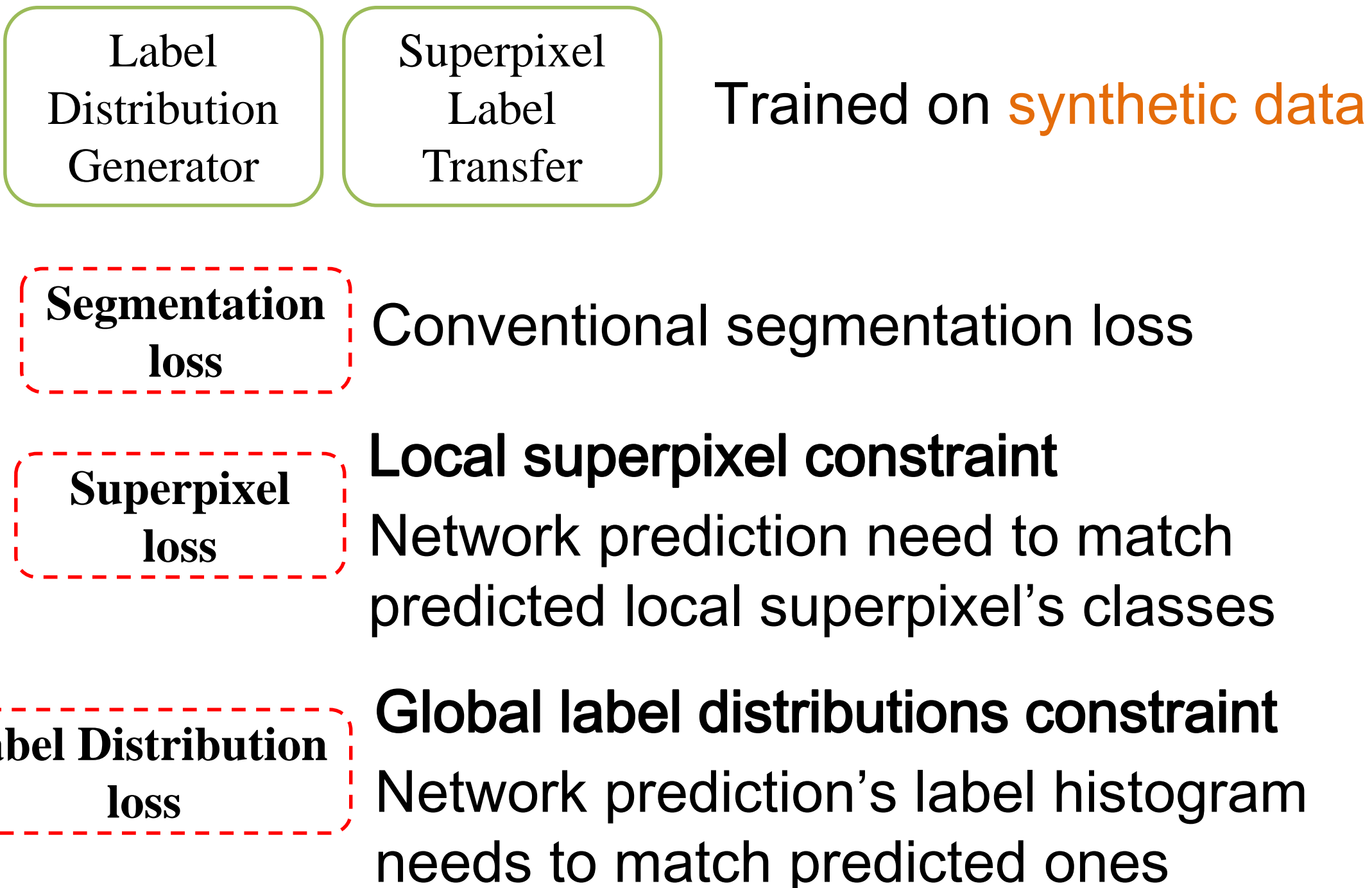How to improve segmentation training **without** accessing real image annotation?

Regularize it with **segmentation posteriors of the target domain**, which are learned from easier tasks

Posterior 1: Global label distributions of images
Constrain segmentation CNN training w/ predicted label histograms of real images

Posterior 2: Local label distributions of landmark superpixels
Constrain segmentation CNN training w/ predicted **chosen confident** superpixel classes of real images

Label Distribution Generator   Superpixel Label Transfer   Trained on synthetic data

Segmentation loss   Conventional segmentation loss

Superpixel loss   **Local superpixel constraint**
Network prediction need to match predicted local superpixel's classes

Label Distribution loss   **Global label distributions constraint**
Network prediction's label histogram needs to match predicted ones
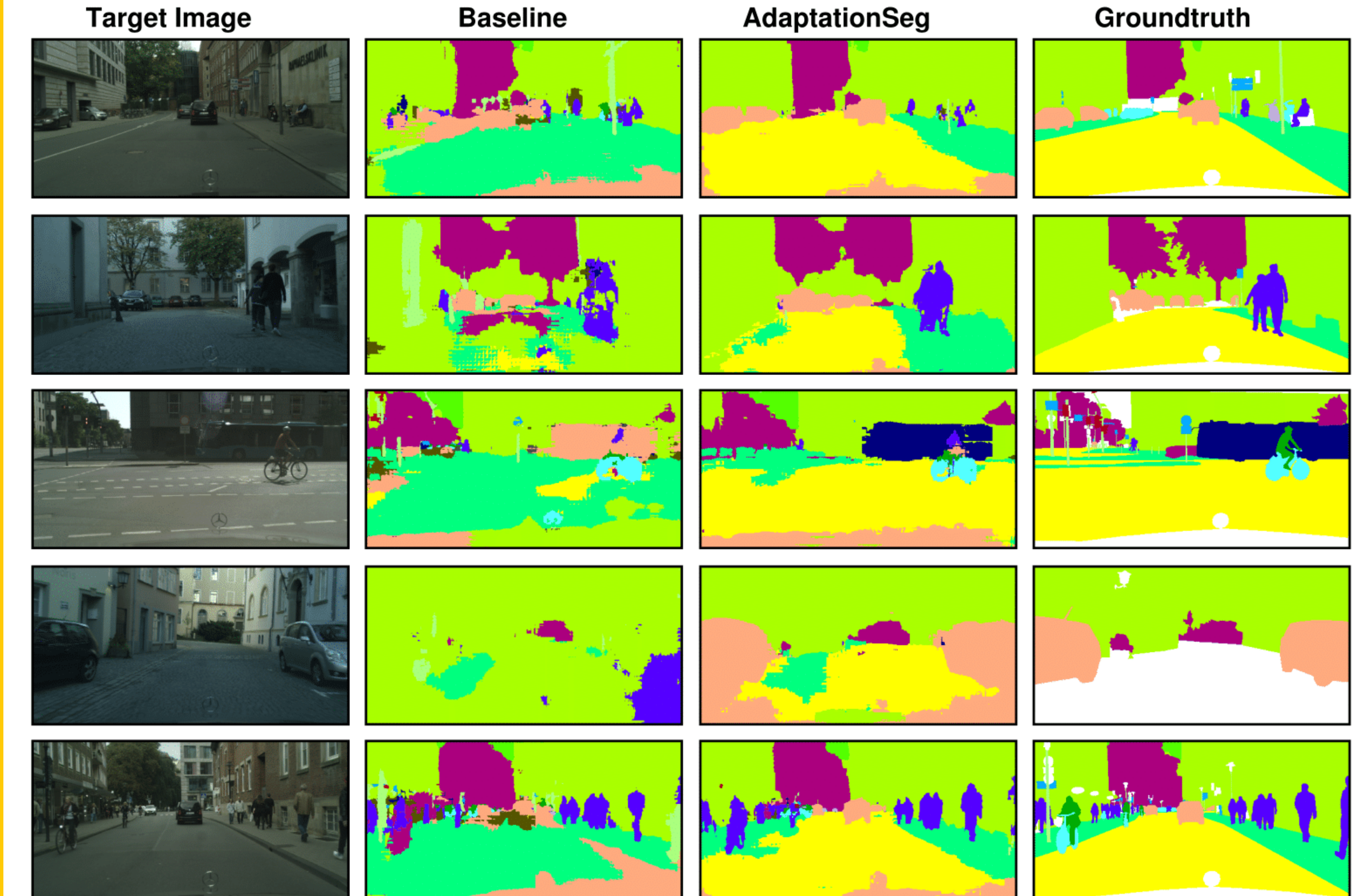
Conclusion

Proposed a segmentation domain adaptation method:

Infer **properties** of target domain in simpler tasks
Regularize segmentation network by the properties
By adaptation, significantly outperforms original network

Qualitative & Quantitative Results

Target Image	Baseline	AdaptationSeg	Groundtruth

											Class-wise IoU							
Method %	IoU	bike	fence	wall	t-sign	pole	mbike	t-light	sky	bus	rider	veg	bldg	car	person	sidewalk	road
NoAdapt [27]	17.4	0.0	0.0	1.2	7.2	15.1	0.1	0.0	66.8	_3.9_	1.5	30.3	29.7	47.3	51.1	17.7	6.4
FCN Wld [27]	_20.2_	_0.6_	_0.0_	**4.4**	**11.7**	_20.3_	_0.2_	_0.1_	_68.7_	3.2	_3.8_	_42.3_	_30.8_	**54.0**	**51.2**	_19.6_	_11.5_
NoAdapt	22.0	**18.0**	0.5	0.8	5.3	21.5	0.5	8.0	_75.6_	4.5	**9.0**	_72.4_	59.6	23.6	_35.1_	11.2	5.6
Ours (I)	_25.5_	16.7	**0.8**	_2.3_	_6.4_	**21.7**	**1.0**	9.9	59.6	_12.1_	7.9	70.2	_67.5_	_32.0_	29.3	_18.1_	_51.9_
SP Lndmk	23.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	**83.1**	**26.1**	0.0	73.1	67.7	41.1	5.8	10.6	60.8
SP	25.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	80.5	22.1	0.0	71.9	69.3	_45.9_	24.6	19.8	**75.0**
Ours (SP)	_28.1_	_10.2_	_0.4_	_0.1_	_2.7_	_8.1_	_0.8_	_3.7_	68.7	21.4	_7.9_	_75.5_	_74.6_	42.9	_47.3_	_23.9_	61.8
Ours (I+SP)	**29.0**	13.1	0.5	0.1	3.0	10.7	0.7	3.7	70.6	20.7	8.2	**76.1**	**74.9**	43.2	47.1	**26.1**	65.2

Code & dataset available at:
https://github.com/YangZhang4065/AdaptationSeg