CrDoCo: Pixel-level Domain Transfer with Cross-Domain Consistency

Yun-Chun Chen[1,2] Yen-Yu Lin[1] Ming-Hsuan Yang[3,4] Jia-Bin Huang[5]

[1]Academia Sinica [2]National Taiwan University [3]UC Merced [4]Google [5]Virginia Tech

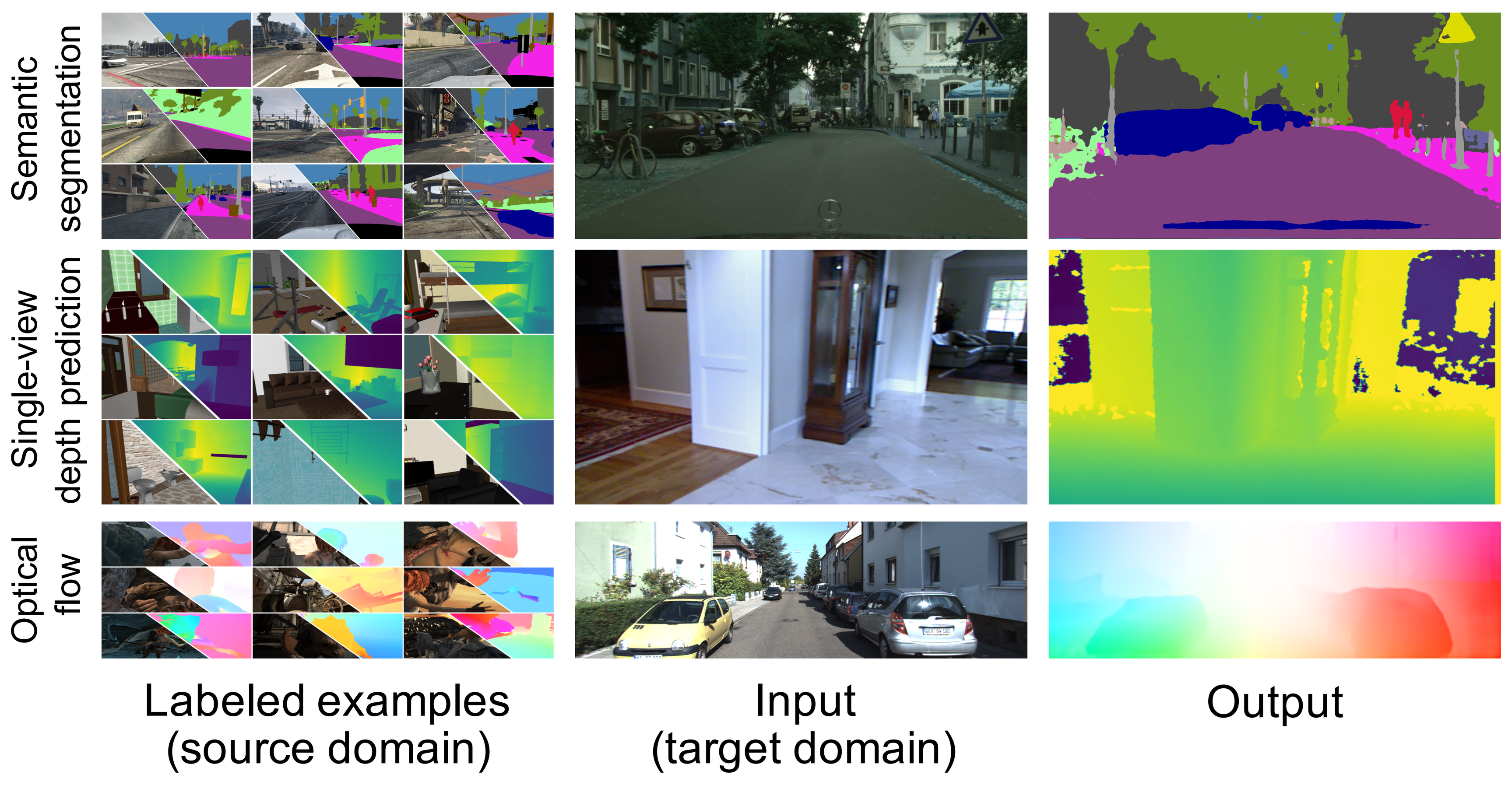
http://bit.ly/CrDoCo

Fork me on GitHub

Unsupervised domain adaptation

Input: A source dataset (labeled) and a target dataset (unlabeled)

Goal: Transfer knowledge from source to target domains

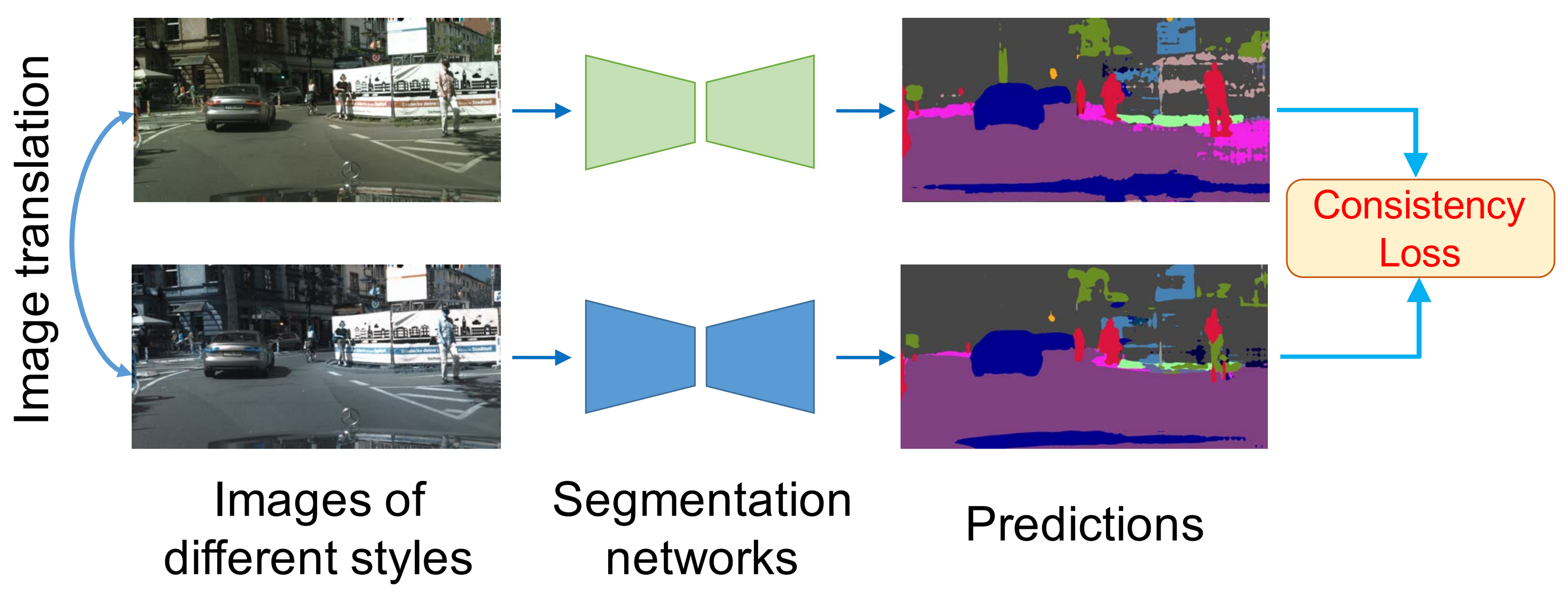


Challenges

- Target domain is unsupervised
- Domain gap between the source and target datasets

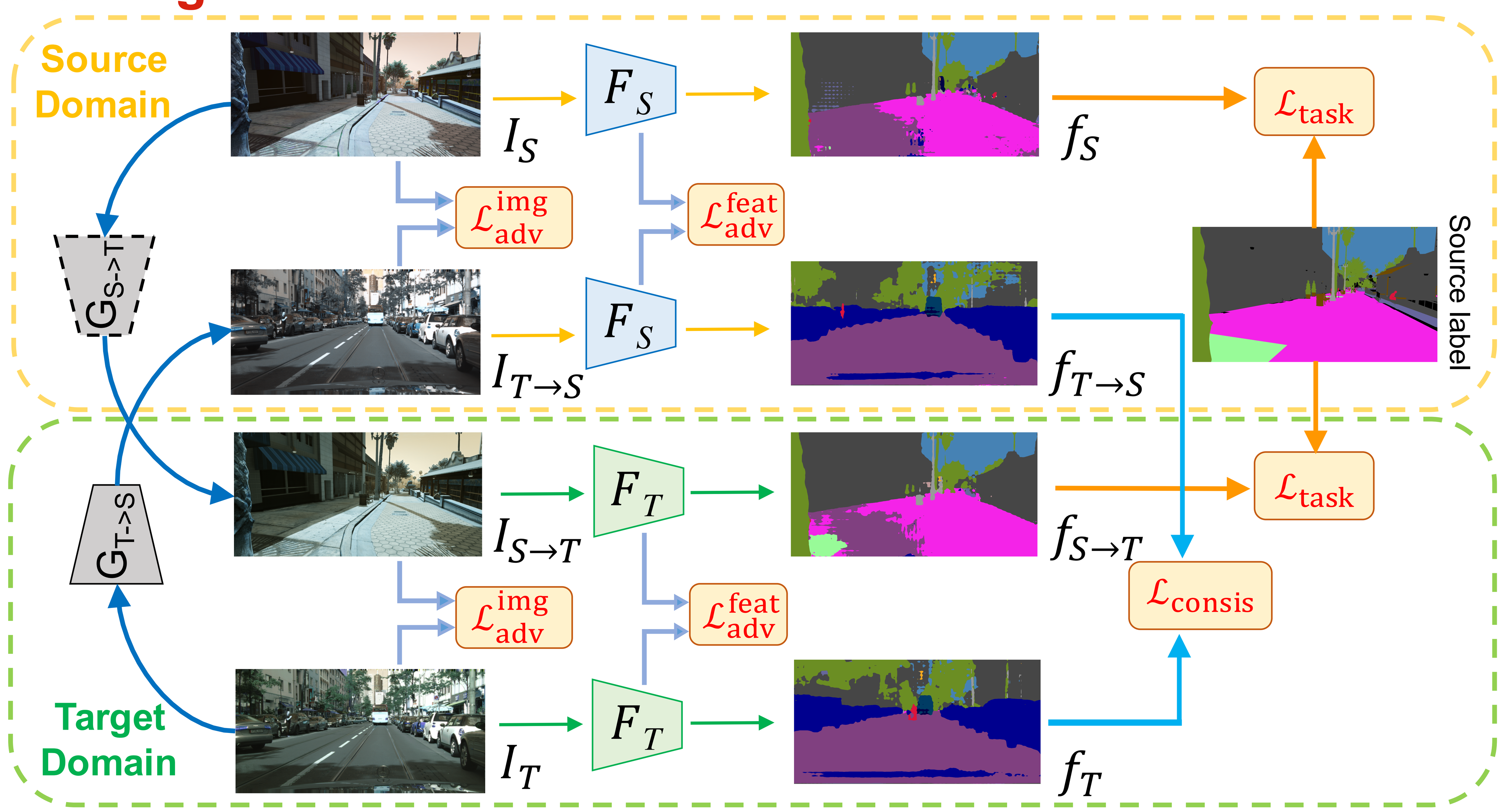
Cross-domain consistency

- Images of different styles should have the same task predictions

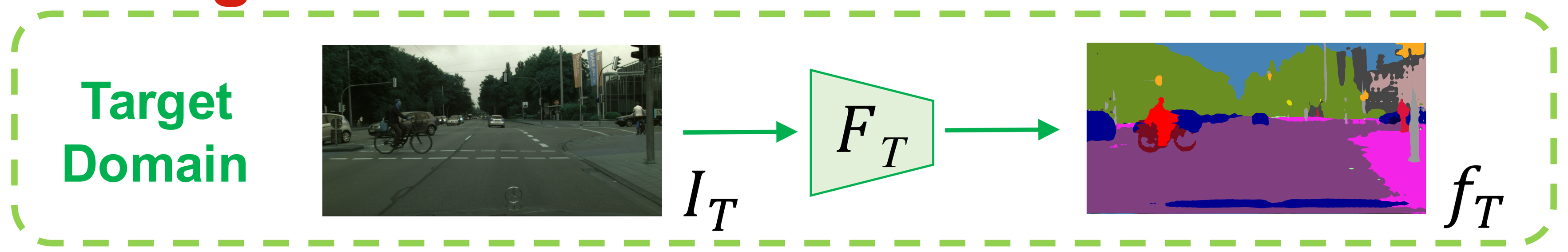


Our approach

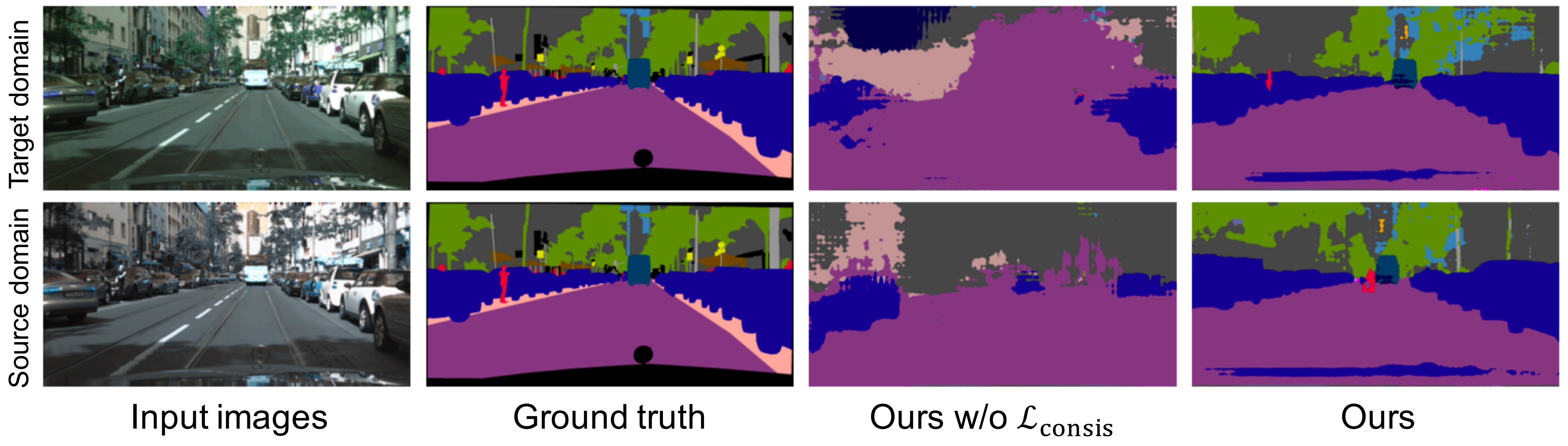
Training



Testing



Cross-domain consistency loss $\mathcal{L}_{\text{consis}}$



Experimental results

A) Synthetic-to-real adaptation

Semantic segmentation

Method	GTA5 → Cityscapes mean IoU	GTA5 → Cityscapes Pixel acc.	SYNTHIA → Cityscapes mean IoU	SYNTHIA → Cityscapes Pixel acc.
Synth.	22.9	71.9	18.5	54.6
DS [Dundar arXiv 18]	38.3	<u>87.2</u>	29.5	<u>76.5</u>
UNIT [Liu NeurIPS 17]	39.1	87.1	28.0	70.8
FCNs ITW [Hoffman arXiv 17]	27.1	-	17.0	-
CyCADA [Hoffman ICML 18]	39.5	82.3	-	-
Ours w/o $\mathcal{L}_{\text{consis}}$	39.4	85.8	<u>29.8</u>	75.3
Ours	**45.1**	**89.2**	**33.4**	**79.5**

Single-view depth prediction

Method	SUNCG → NYUv2 Abs. Rel. ↓	Sq. Rel. ↓	RMSE ↓
Synth.	0.304	0.394	1.024
Baseline (train set mean)	0.439	0.641	1.148
T^2Net [Zheng ECCV 18]	0.257	<u>0.281</u>	0.915
Ours w/o $\mathcal{L}_{\text{consis}}$	<u>0.254</u>	0.283	<u>0.911</u>
Ours	**0.233**	**0.272**	**0.898**

Optical flow estimation

Method	MPI Sintel → KITTI 2012 AEPE *train*	AEPE *test*	F1-Noc *test*	MPI Sintel → KITTI 2015 AEPE *train*	F1-all *train*	F1-all *test*
FlowNet2 [Ilg CVPR 17]	4.09	-	-	<u>10.06</u>	<u>30.37%</u>	-
PWC-Net [Sun CVPR 18]	<u>4.14</u>	<u>4.22</u>	**8.10%**	10.35	33.67%	-
Ours w/o $\mathcal{L}_{\text{consis}}$	4.16	4.92	13.52%	10.76	34.01%	<u>36.43%</u>
Ours	**2.19**	**3.16**	<u>8.57%</u>	**8.02**	**23.14%**	**25.83%**

B) Cross-city adaptation

Method	Cityscapes → Cross-city Rome	Rio	Tokyo	Taipei
Cross-City [Chen ICCV 17]	42.9	42.5	42.8	39.6
CBST [Zou ECCV 18]	<u>53.6</u>	**52.2**	<u>48.8</u>	**50.3**
AdaptSegNet [Tsai CVPR 18]	52.2	49.5	46.9	47.5
Ours w/o $\mathcal{L}_{\text{consis}}$	51.0	48.9	45.9	46.8
Ours	**55.1**	<u>50.4</u>	**51.2**	<u>47.9</u>

C) Visual results