Category-Adaptive for Semantic Segmentation Domain Adaptation

Zhiming Wang, Yantian Luo, Danlan Huang, Ning Ge, Jianhua Lu

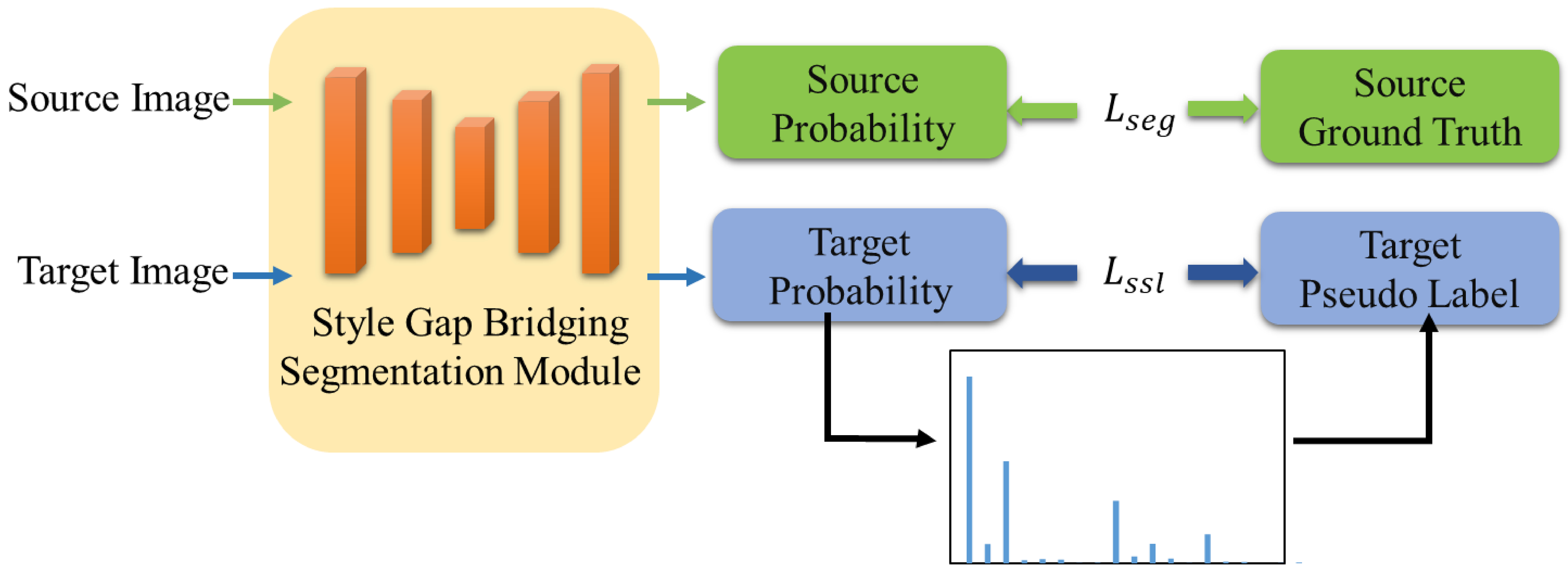
Department of Electronic Engineering, Tsinghua University, Beijing 100084, P.R. China

Beijing National Research Center for Information Science and Technology

icassp 2022 Singapore

Paper ID:8856

Cross-Domain Adaptation for Semantic Segmentation

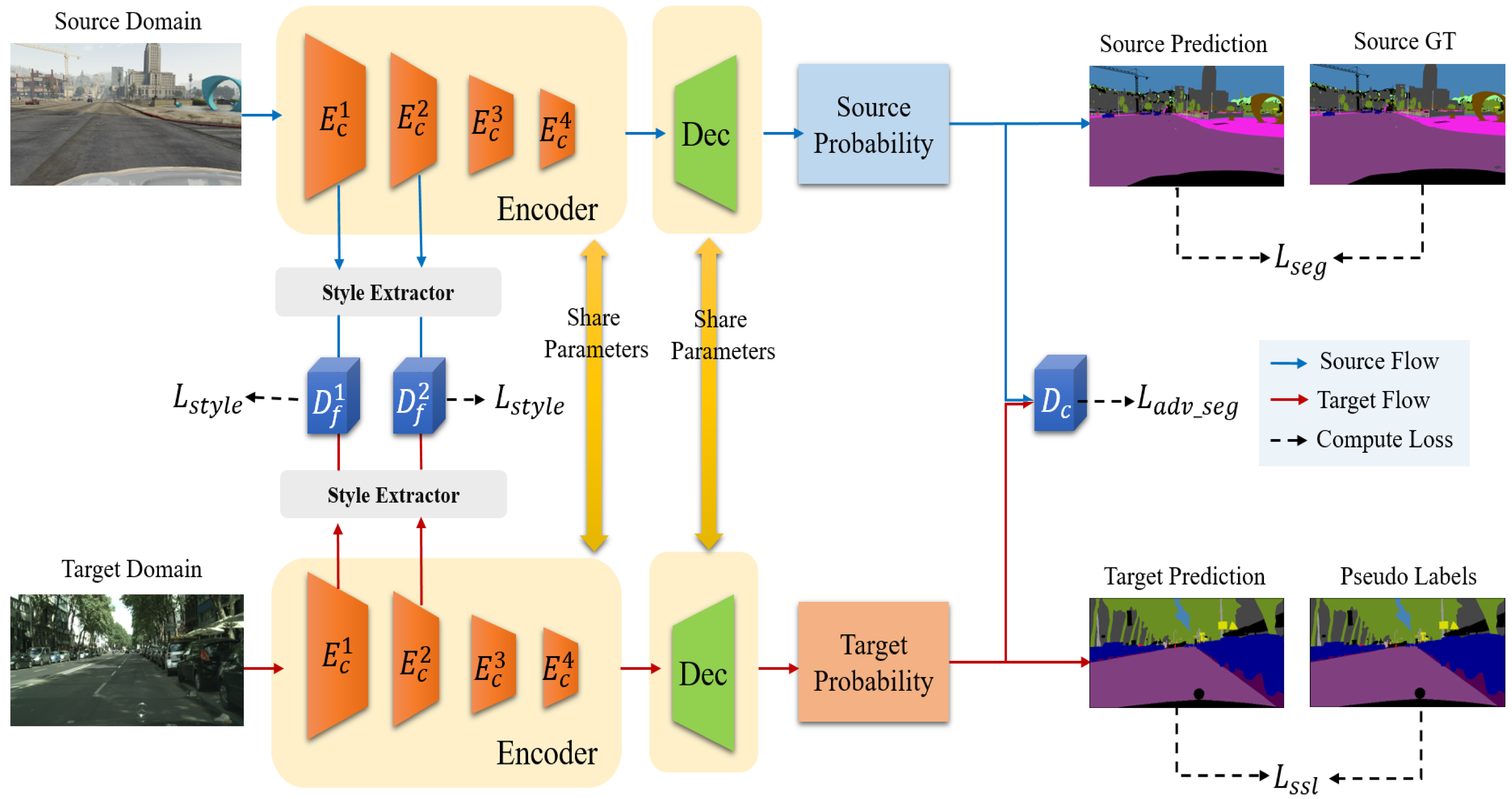


Category Adaptive Thresholds

Our contributions:

- Adversarial learning is introduced for the style gap bridging mechanism.
- A new category-adaptive threshold mechanism is proposed to choose pseudo labels for self-supervised learning.

Architecture of Proposed Method



Loss functions:

1. Segmentation Loss $\mathcal{L}_{seg}$
2. Domain Adaptation Loss $\mathcal{L}_{adv\_seg}$
3. Style loss $\mathcal{L}_{style}$
4. Self-supervised Loss $\mathcal{L}_{ssl}$

Pseudo Labeling Mechanism

1. Define the category centroid of class c:

$$f^c = \frac{1}{|\mathcal{P}^c|} \sum_{|\mathcal{X}_t|} \sum_{h=1}^{H_t} \sum_{w=1}^{W_t} \mathbb{1}_{[c=\arg\max_{c'} P_t^{hwc'}]} P_t^{hw}$$

2. Define pseudo labels m_t^{hwc} based on entropy comparison and given threshold $\triangle$:

$$m_t^{hwc} = \mathbb{1}_{[H(P_t^{hw})<H(f^c)-\Delta,\, c=\arg\max_{c'} P_t^{hwc'}]}$$

3. self-supervised learning based on pseudo labels:

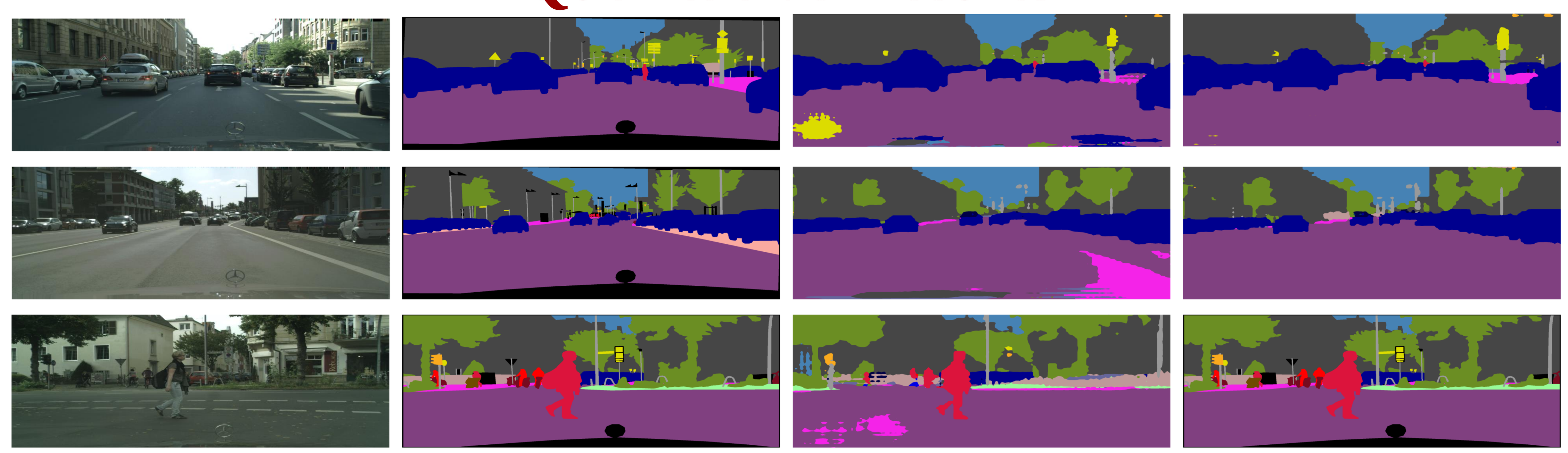
$$\mathcal{L}_{ssl} = -\frac{1}{H_t W_t} \sum_{h=1}^{H_t} \sum_{w=1}^{W_t} \sum_{c=1}^{C} m_t^{hwc} \hat{y}_t^{hwc} \log P_t^{hwc}$$

Quantitative Results

GTA5 $\rightarrow$ Cityscapes

Method	road	sidewalk	building	wall	fence	pole	t-light	t-sign	vegetation	terrain	sky	person	rider	car	truck	bus	train	motorbike	bicycle	mIoU
CBST[4]	89.6	**58.9**	78.5	33.0	22.3	**41.4**	**48.2**	39.2	83.6	24.3	65.4	49.3	20.2	83.3	39.0	48.6	**12.5**	20.3	35.3	47.0
Cycada [19]	86.7	35.6	80.1	19.8	17.5	38.0	39.9	**41.5**	82.7	27.9	73.6	**64.9**	19	65.0	12.0	28.6	4.5	31.1	42.0	42.7
ADVENT [6]	87.6	21.4	82.0	34.8	26.2	28.5	35.6	23.0	84.5	35.1	76.2	58.6	30.7	84.8	34.2	43.4	0.4	28.4	35.2	44.8
DCAN [20]	85.0	30.8	81.3	25.8	21.2	22.2	25.4	26.6	83.4	36.7	76.2	58.9	24.9	80.7	29.5	42.9	2.5	26.9	11.6	41.7
CLAN [21]	87.0	27.1	79.6	27.3	23.3	28.3	35.5	24.2	83.6	27.4	74.2	58.6	28.0	76.2	33.1	36.7	6.7	**31.9**	31.4	43.2
BDL [5]	91.0	44.7	84.2	34.6	**27.6**	30.2	36.0	36.0	**85.0**	**43.6**	83.0	58.6	31.6	83.3	35.3	49.7	3.3	28.8	35.6	48.5
Ours	**91.7**	51.1	**85.0**	**38.7**	26.7	32.1	38.1	34.6	84.3	38.6	**84.9**	60.7	**32.8**	**85.2**	**41.9**	**49.8**	2.8	28.5	**45.0**	**50.2**

Qualitative Results



(a) Image (b) GT (c) BDL (d) Ours

Style Gap Bridging Comparison

style gap bridging mechanism	style modeling	mIoU
MSE	Gram matrix	44.7
	mean & std	45.1
adversarial learning	mean (Ours)	45.5

Ablation Studies

GTA5 $\rightarrow$ Cityscapes

model	mIoU
original	44.6
original + adv	45.5
original + adv + SSL once	48.5
original + adv + SSL twice	50.2