Perturbed and Strict Mean Teachers for Semi-supervised Semantic Segmentation

Yuyuan Liu[1], Yu Tian[1], Yuanhong Chen[1], Fengbei Liu[1], Vasileios Belagiannis[2], Gustavo Carneiro[1]

Australian Institute for Machine Learning [1]   Universitat Ulm, Germany [2]

AUSTRALIAN INSTITUTE FOR MACHINE LEARNING

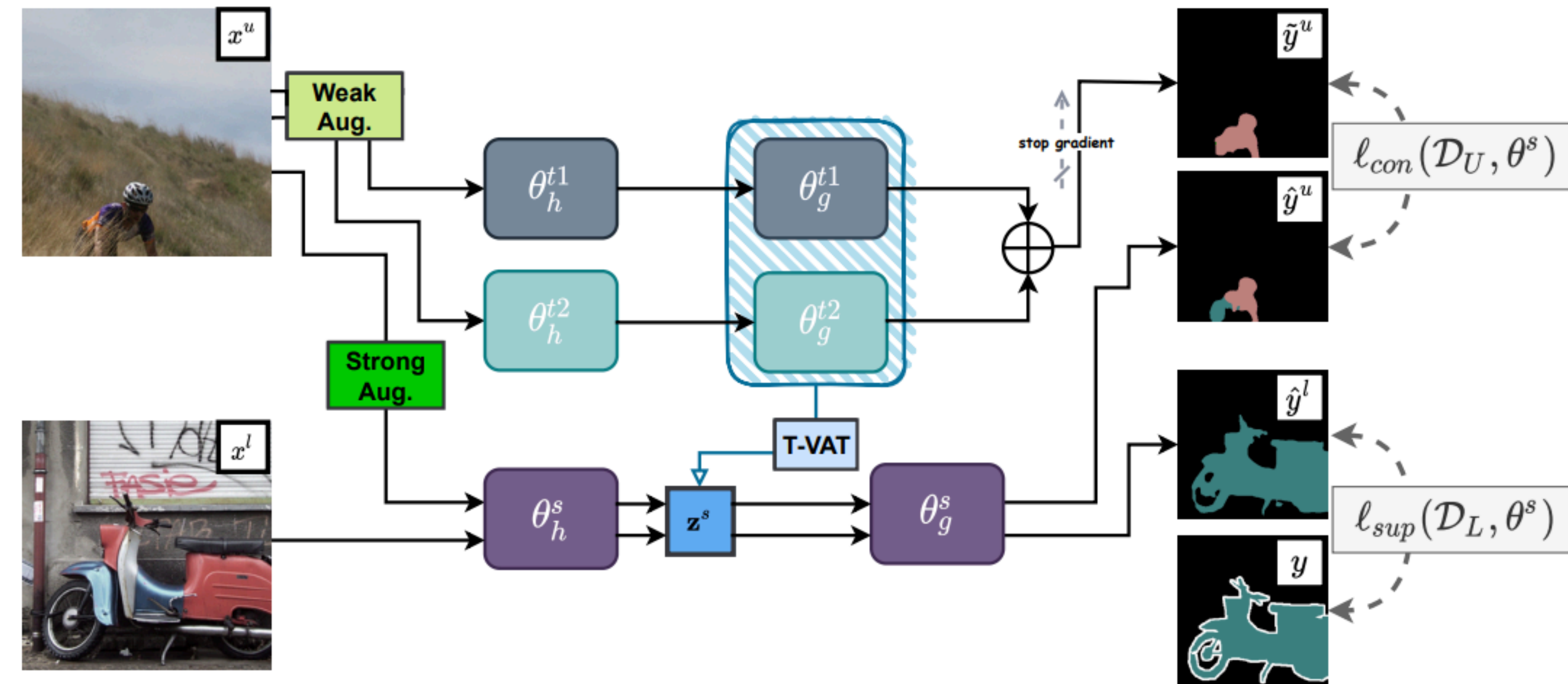THE UNIVERSITY of ADELAIDE

CVPR JUNE 19-24 2022 NEW ORLEANS · LOUISIANA

Motivation

✦ Inaccurate pseudo-labels hinder the consistency regularisation.

✦ Vanilla Mean Teacher (MT) yields insufficient perturbation for semi-supervised segmentation.

✦ MSE loss leads the gradient to vanish, while CE easily overfits the noise signal (leads to confirmation bias).

Contribution
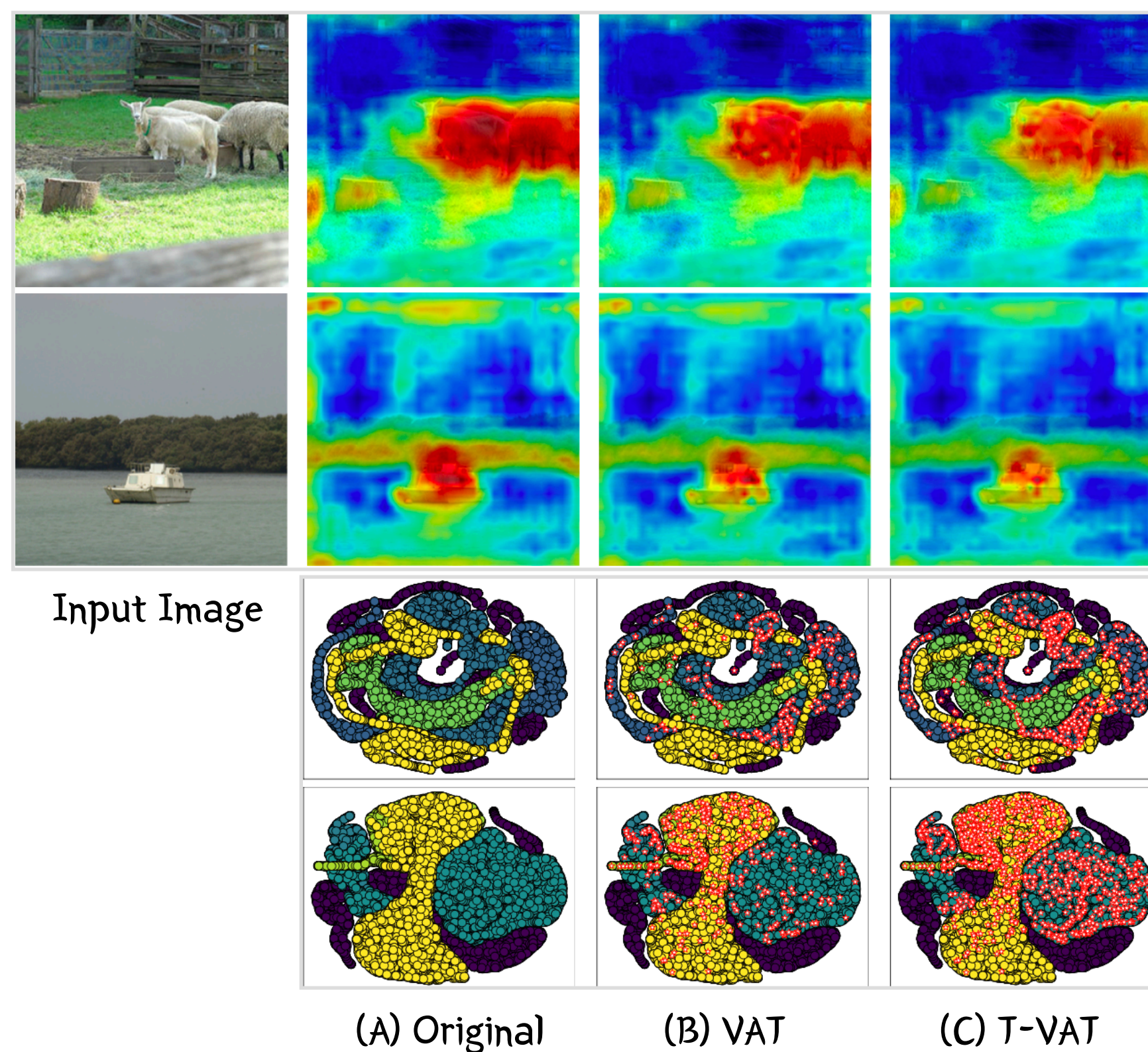
✦ We include a new auxiliary teacher to improve the pseudo-label quality further.

✦ We introduce a novel teacher-based virtual adversarial training (T-VAT) noise to enhance consistency regularisation.

✦ We assign the teachers' ensemble confidence (conf-CE) to alleviate confirmation bias for learning the unlabelled data.

✦ Our model outperforms other SOTA approaches in both Pascal VOC12 and Cityscapes datasets.
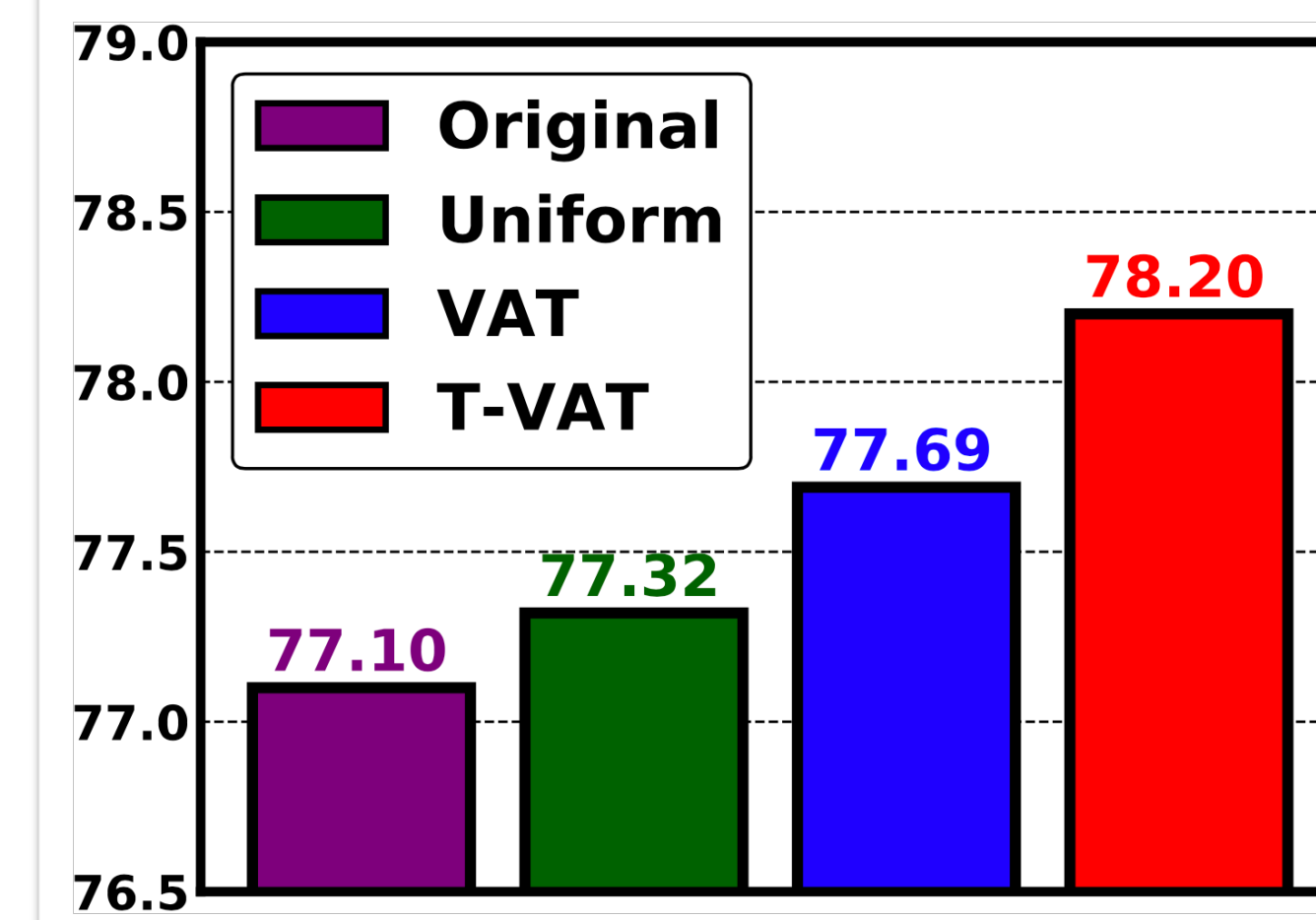
Methodology

Effectiveness of T-VAT

a. visualisation

Input Image

(A) Original   (B) VAT   (C) T-VAT

b. Feature perturb. improvements

★ In Attention Visualisation, T-VAT confuses the student model the most.

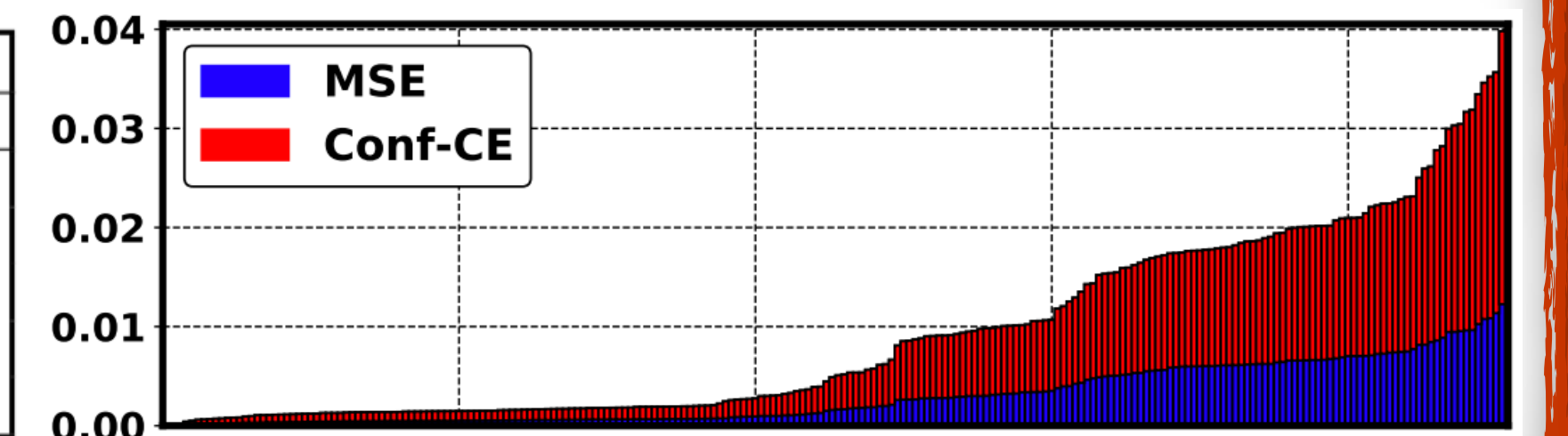★ In TSNE Visualisation, T-VAT leads to better consistency regularisation.
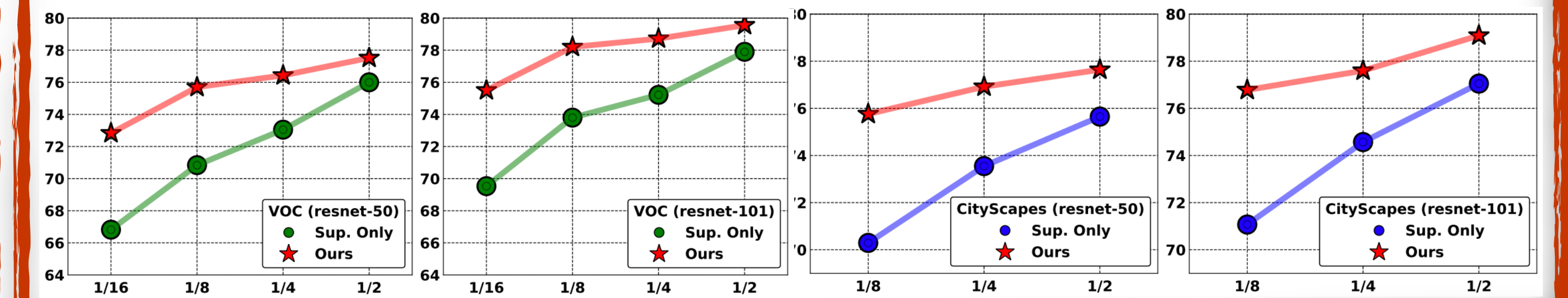
Experiments   (measured by mIoU)

a. ablation study

MT	conf-CE	T-VAT	AT	Backbone	
				ResNet-50	ResNet-101
✓				71.49	73.50
✓	✓			73.79	76.39
✓	✓	✓		74.87	77.36
✓	✓		✓	74.92	77.47
✓	✓	✓	✓	**75.70**	**78.20**

b. The different gradients from losses

MSE
Conf-CE

c. comparing with sup. Baselines

VOC (resnet-50)   VOC (resnet-101)   CityScapes (resnet-50)   CityScapes (resnet-101)
● Sup. Only ★ Ours

d. comparing with other SOTA on VOC12

Method	ResNet-50				ResNet-101			
	1/16(662)	1/8(1323)	1/4(2646)	1/2(5291)	1/16(662)	1/8(1323)	1/4(2646)	1/2(5291)
MT* [37]	66.77	70.78	73.22	75.41	70.59	73.20	76.62	77.61
French et al.* [14]	68.90	70.70	72.46	74.49	72.56	72.69	74.25	75.89
CCT* [34]	65.22	70.87	73.43	74.75	67.94	73.00	76.17	77.56
GCT* [21]	64.05	70.47	73.45	75.20	69.77	73.30	75.25	77.14
ECS [31]	-	67.38	70.70	72.89	-	-	-	-
CPS [9]	71.98	73.67	74.90	76.15	74.48	76.44	77.68	78.64
CAC [22]	70.10	72.40	74.00	-	72.40	74.60	76.30	-
Ours	**72.83**	**75.70**	**76.43**	**77.88**	**75.50**	**78.20**	**78.72**	**79.76**

e. comparing with other SOTA on City

Method	Backbone	1/8	1/4	1/2
ECS [31]	ResNet50	67.38	70.70	72.89
CAC [22]	ResNet50	69.70	72.70	-
Ours	ResNet50	74.37	75.15	76.02
Ours (sliding eval.)	ResNet50	75.76	76.92	77.64
	ResNet101	76.89	77.60	79.09
GCT [21]†	ResNet50	71.33	75.30	77.09
CPS [9]†	ResNet50	76.61	77.83	78.77
Ours†	ResNet50	77.12	78.38	79.22

f. additional data exp.

Baseline   Ours
ResNet50   ResNet101
78.94 80.13   79.95 81.02