



# Perturbed and Strict Mean Teachers for Semi-supervised Semantic Segmentation

Yuyuan Liu<sup>1</sup>, Yu Tian<sup>1</sup>, Yuanhong Chen<sup>1</sup>, Fengbei Liu<sup>1</sup>, Vasileios Belagiannis<sup>2</sup>, Gustavo Carneiro<sup>1</sup>

Australian Institute for Machine Learning <sup>1</sup> Otto von Guericke University <sup>2</sup>

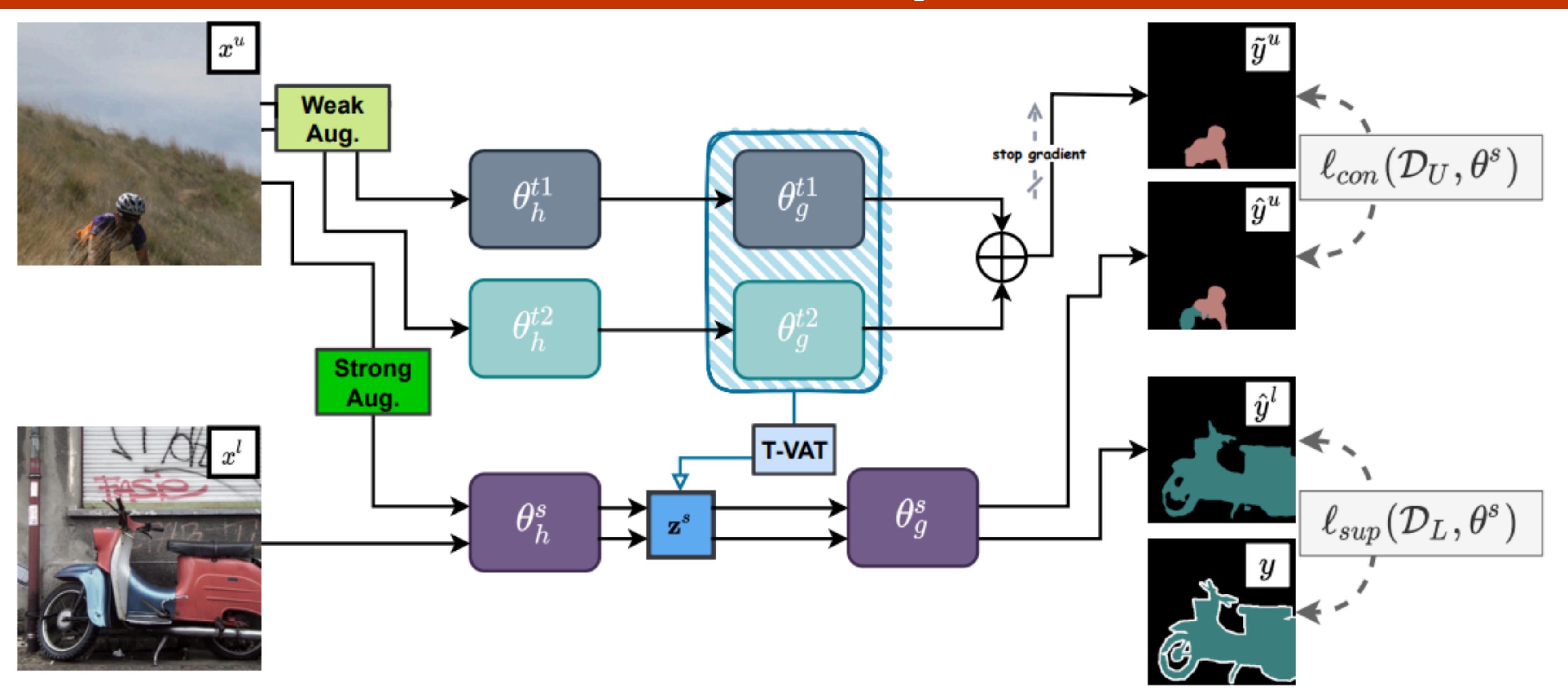
## Motivation

- Inaccurate pseudo-labels turn down 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

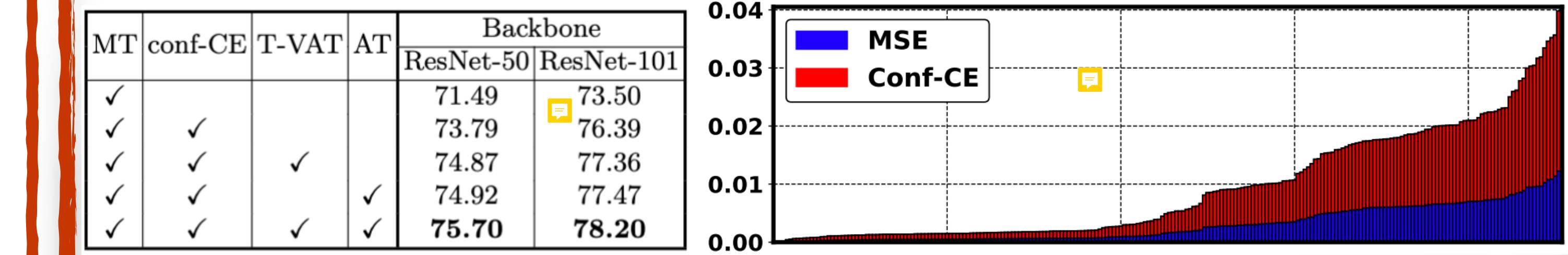
- We expand a new auxiliary teacher to stabilise the pseudo-label quality further.
- We introduce a novel teacher-based virtual adversarial training (T-VAT) noise to enhance consistency regularisation.
- We assign the teachers' ensemble confidence (conf-CE) to alleviate confirmation bias for learning the unlabelled data.
- Our model outperforms other SOTA approaches in both Pascal VOC12 and Cityscapes datasets.

## Methodology

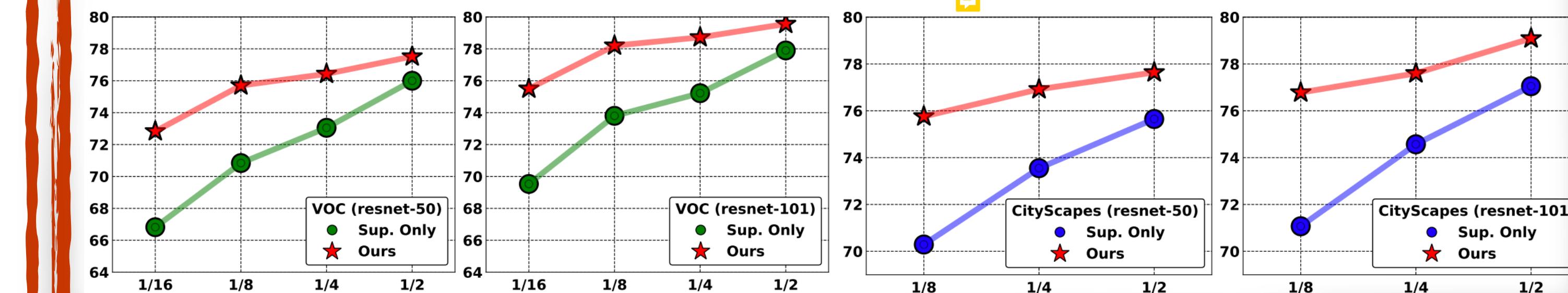


## Experiments

### a. ablation study



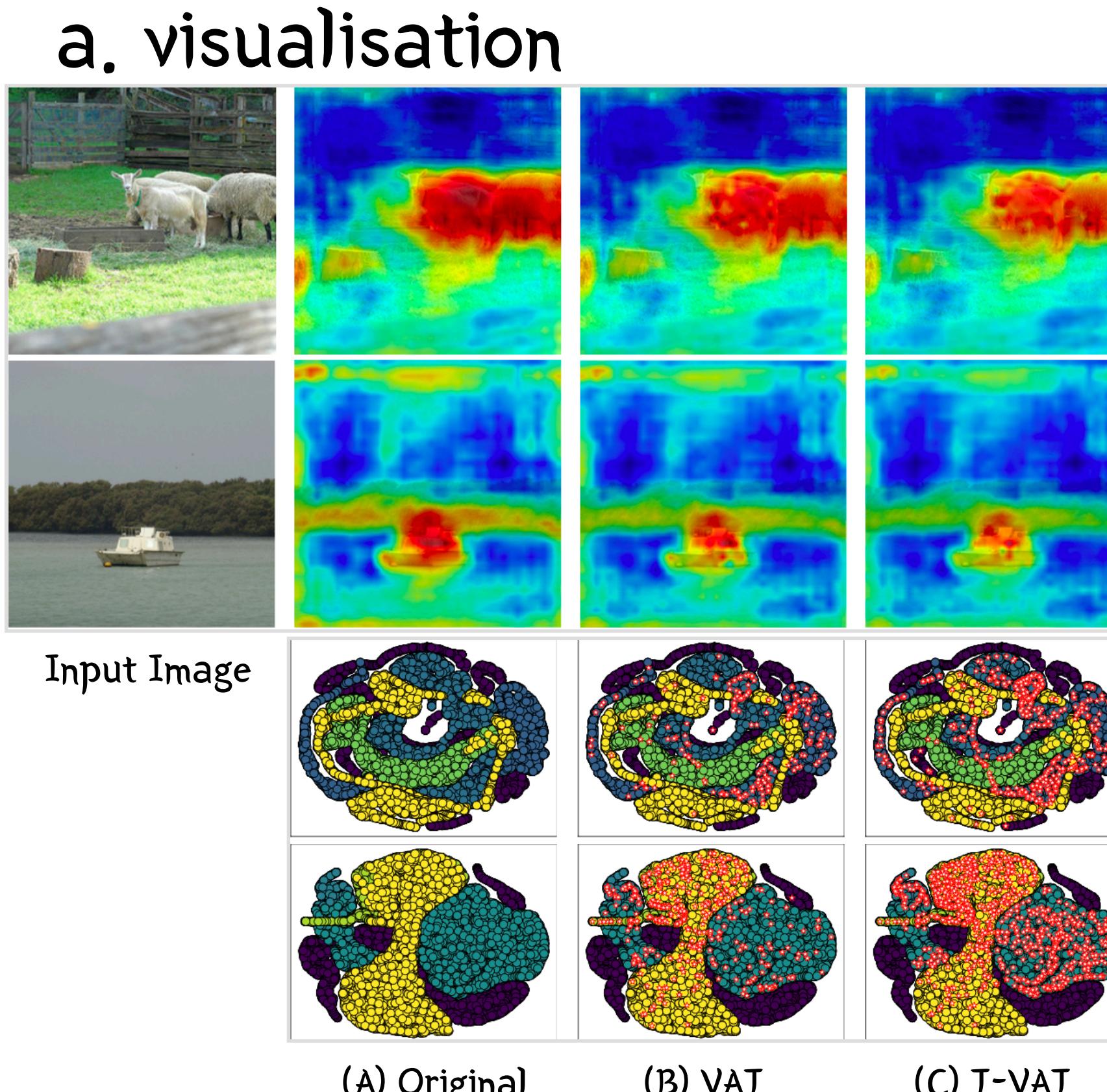
### b. comparing with other sup. baselines



### c. 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

## Effectiveness of T-VAT



- In Attention Visualisation, T-VAT confuses the student model the most.
- In TSNE Visualisation, T-VAT leads to better consistency regularisation.

### d. comparing with other SOTA on City

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

### e. additional data exp.

