A Strong Baseline for Generalized Few-Shot Semantic Segmentation

WED-AM-289

Sina Hajimiri

Malik Boudiaf

Ismail Ben Ayed

Jose Dolz

ÉTS Montreal, Canada

ÉTS
Engineering for Industry
ÉCOLE DE TECHNOLOGIE SUPÉRIEURE
Université du Québec

JUNE 18-22, 2023
CVPR
VANCOUVER, CANADA

Few-shot segmentation (FSS)

$\mathcal{Y}_{\text{base}} \cap \mathcal{Y}_{\text{novel}} = \varnothing$

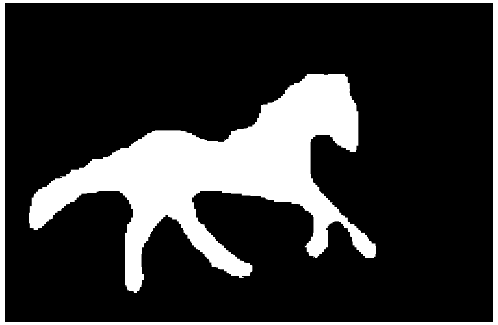
Training: $\mathcal{Y}_{\text{base}} = \{\text{car},\ \text{cat},\ \text{person}\}$

Testing: $\text{horse} \sim \mathcal{Y}_{\text{novel}}$

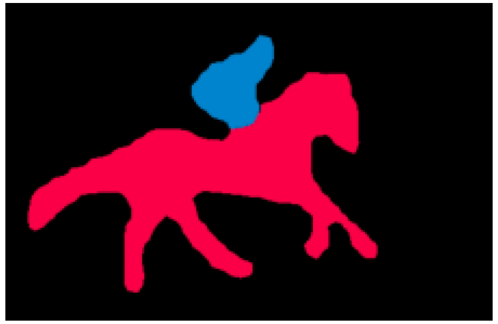
Support set

Query

FSS



GFSS



Standard FSS only segments the novel class,
disregarding the previously learned base classes

Motivation

- Prevent knowledge loss
- No knowledge of novel classes beforehand
- No need for base classes' labels in the support set
- Novel classes performance

Contributions

Presenting *Distilled Information Maximization* (DIaM)

Consistent SOTA on PASCAL-5^i and COCO-20^i

Outperforming current SOTA in a more challenging benchmark

Objective breakdown

Mutual Information

Between classifier's input (X) and output (P)

$-\mathrm{H}(P|X)$

Conditional entropy

$\mathrm{H}(P)$

Marginal entropy

$-\sum_{i=1}^{|\mathbb{S}|} \mathrm{CE}(\mathbf{y}_i; \pi_{\mathbb{S}}(\mathbf{p}_i))$

Cross-entropy on the support images

$-\mathrm{H}(\mathbf{p}_Q)$

Shannon entropy on the query image

$-\mathrm{KL}(\widehat{\mathbf{p}} \,||\, \mathbf{\Pi})$

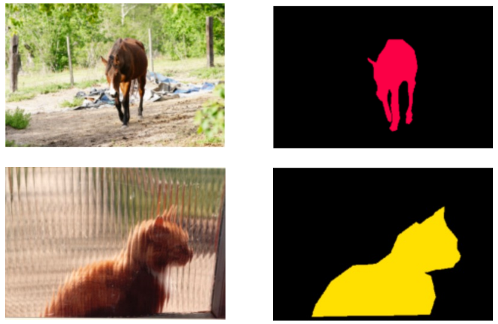
Between model's marginal and a self-estimated distribution

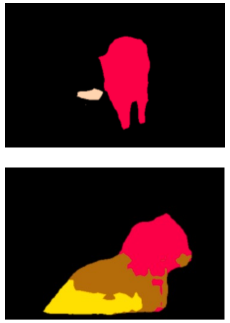
... kite horse bear $\in \mathcal{Y}_{\text{novel}}$

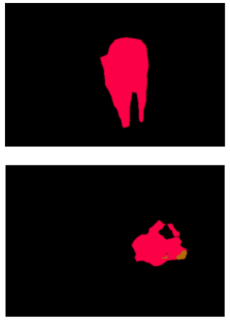
... cat $\in \mathcal{Y}_{\text{base}}$

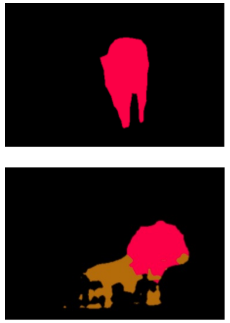
Query

Ground truth









... kite horse bear $\in \mathcal{Y}_{\text{novel}}$

... cat $\in \mathcal{Y}_{\text{base}}$

(negative) Mutual Information

Between classifier's input (X) and output (P)

$$\sum_{i=1}^{|\mathbb{S}|} \mathrm{CE}(\mathbf{y}_i; \pi_{\mathbb{S}}(\mathbf{p}_i))$$

Cross-entropy on the support images

$$\mathrm{H}(\mathbf{p}_Q)$$

Shannon entropy on the query image

$$\mathrm{KL}(\widehat{\mathbf{p}} \,||\, \mathbf{\Pi})$$

Between model's marginal and a self-estimated distribution

Knowledge Distillation

Between old trained model and current model

$$\mathrm{KL}(\pi_{\text{new2old}}(\mathbf{p}_Q) \,||\, \mathbf{p}_Q^{\text{old}})$$

Between current predictions and old snapshot's

Query Ground truth

Experiments

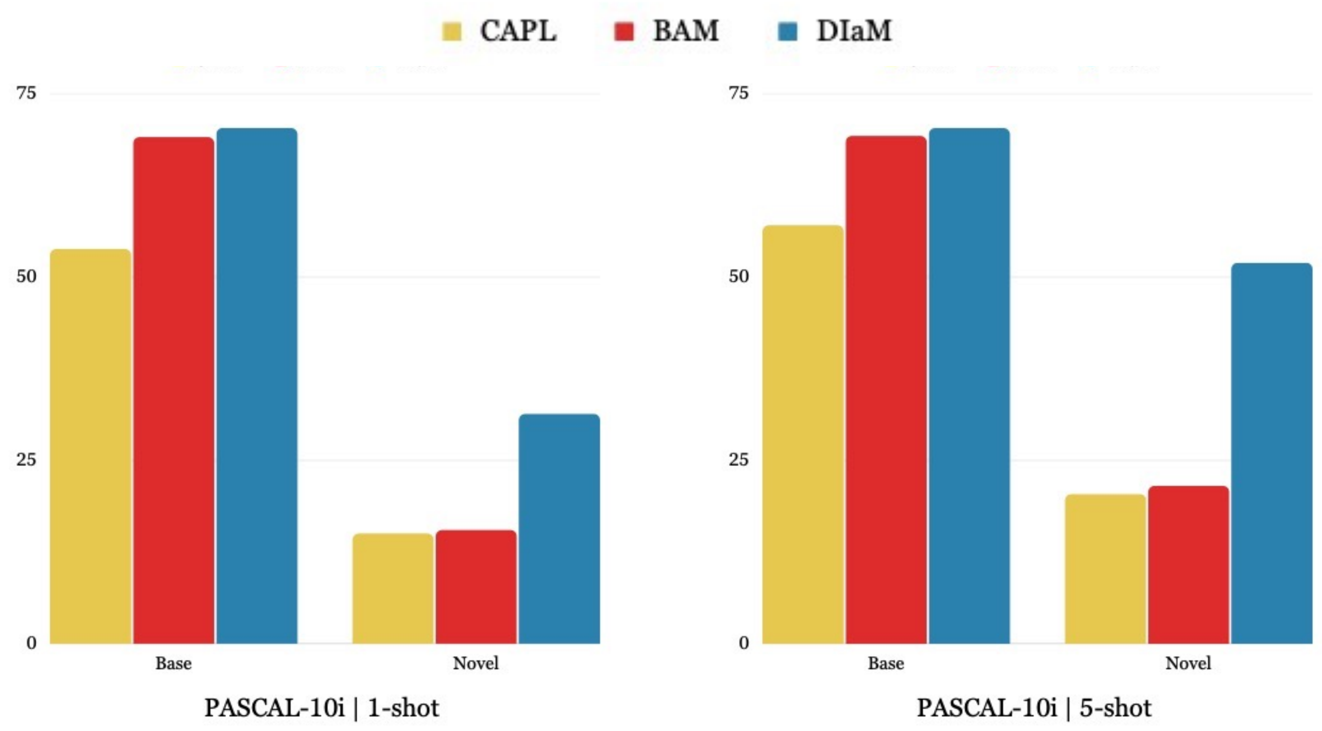
Method		PASCAL-5^i				
	1-Shot			5-Shot		
Method	Base	Novel	Mean	Base	Novel	Mean
CAPL (Tian et al., 2022)	64.80	17.46	41.13	65.43	24.43	44.93
BAM (Lang et al., 2022)	**71.60**	27.49	49.55	**71.60**	28.96	50.28
DIaM (Ours)	70.89	**35.11**	**53.00**	70.85	**55.31**	**63.08**

	COCO-20^i					
	Base	Novel	Mean	Base	Novel	Mean
CAPL (Tian et al., 2022)	43.21	7.21	25.21	43.71	11.00	27.36
BAM (Lang et al., 2022)	**49.84**	14.16	32.00	**49.85**	16.63	33.24
DIaM (Ours)	48.28	**17.22**	**32.75**	48.37	**28.73**	**38.55**

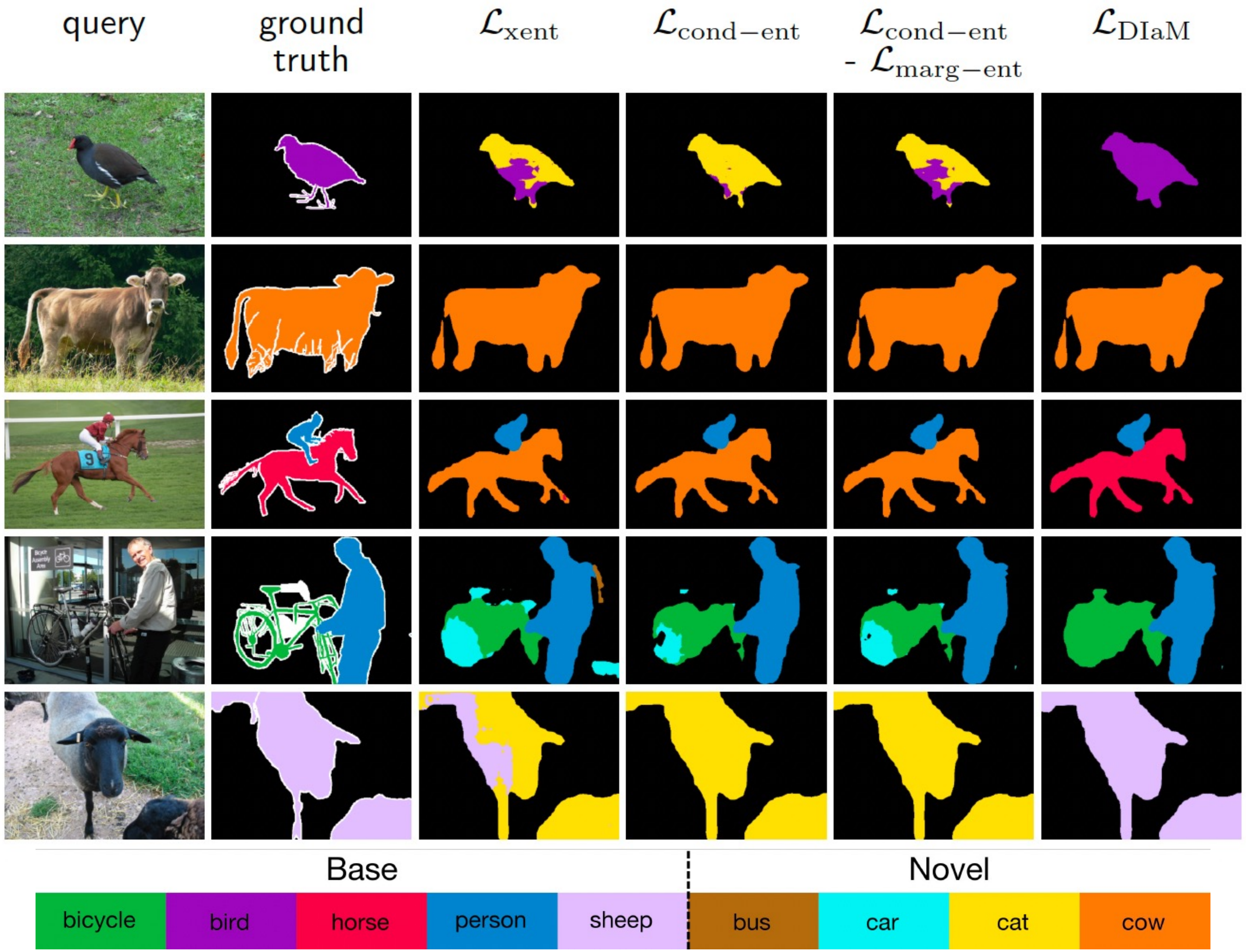
- Keep high performance on base classes
- Outperform on novel classes
- Better leverage information as more shots are given

Experiments (cont.)

- Equal number of base and novel classes



Visual examples



Conclusion

Main takeaways

- Standard FSS can be made more practical → GFSS
- Unrealistic assumptions in prior works (e.g., knowing novel classes beforehand)
- A baseline to eliminate these assumptions

Future directions

- Better estimation of Π
- Avoiding ambiguity of the *background* class

Learn more

- Camera-ready: https://arxiv.org/abs/2211.14126
- Code: https://github.com/sinahmr/DIaM