


Improving Zero-shot Generalization of Learned Prompts via Unsupervised Knowledge Distillation

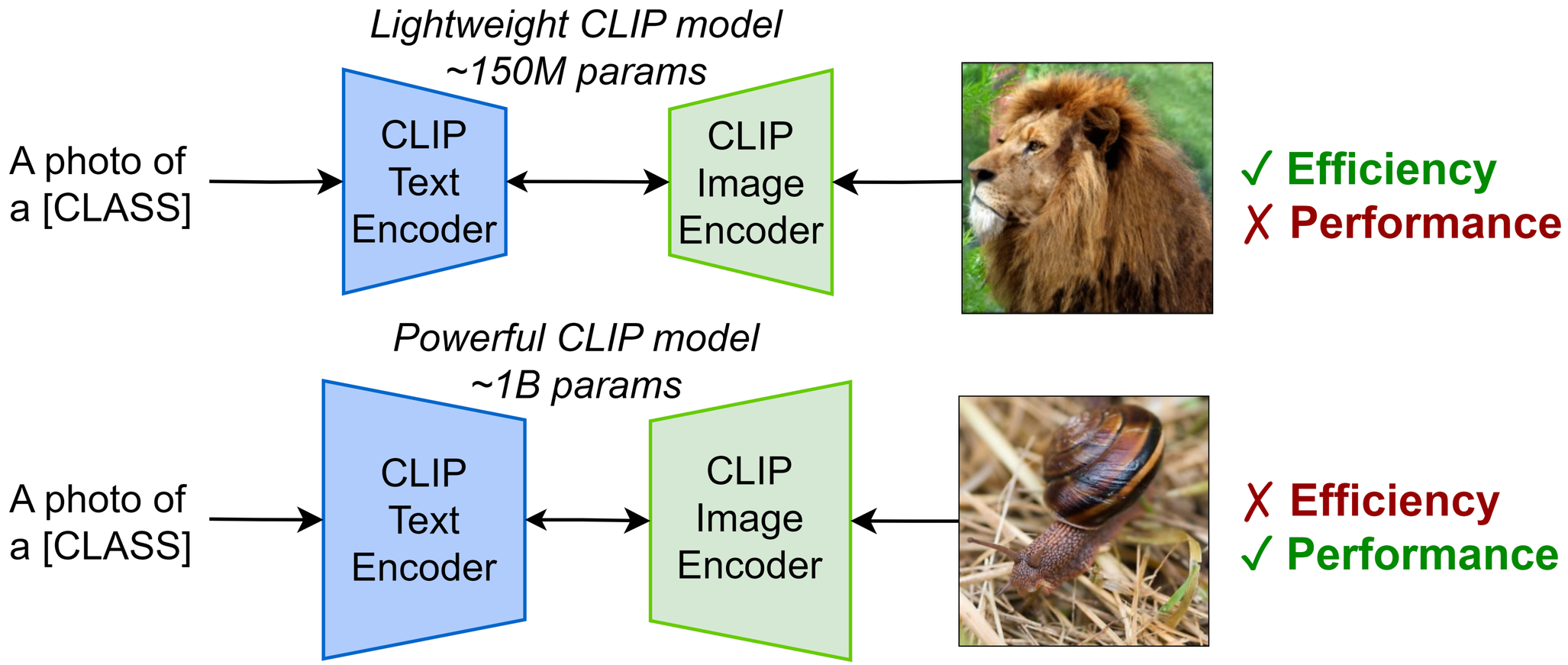
*Marco Mistretta, *Alberto Baldrati, Marco Bertini, Andrew D. Bagdanov

arXiv GitHub

Presented by Andrew D. Bagdanov

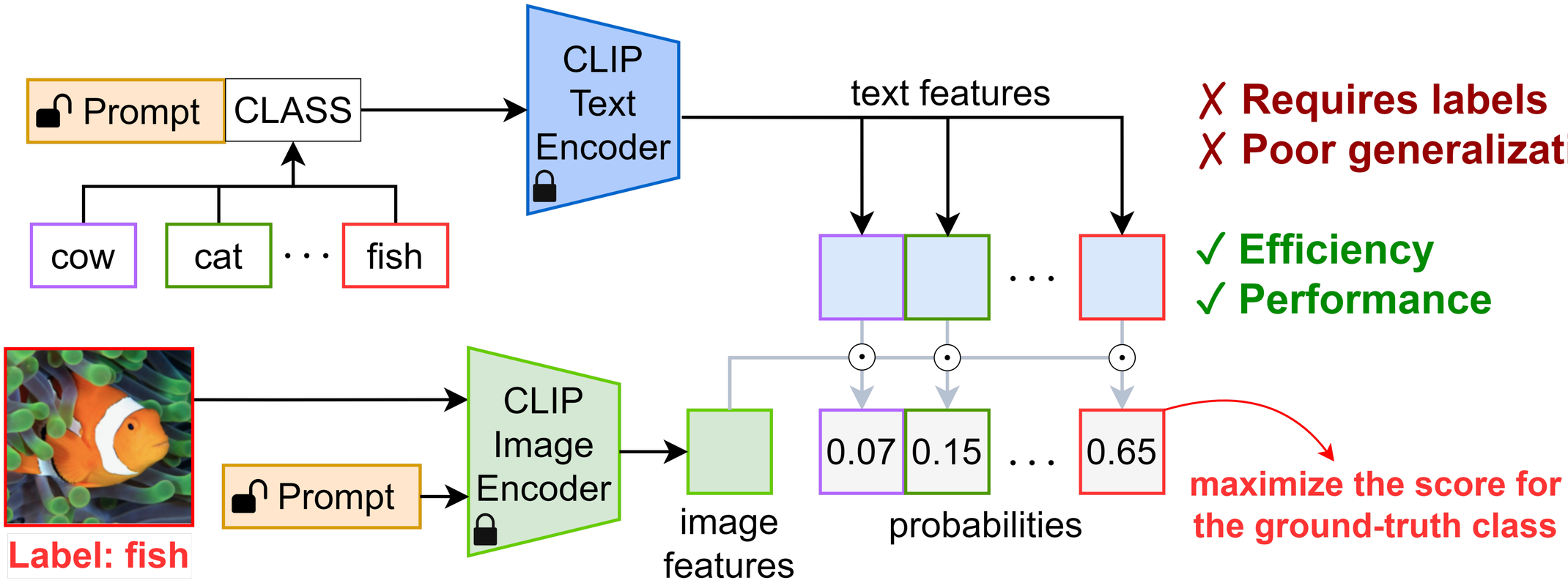
Motivation

Larger CLIPs are powerful but are not efficient. Fine-tuning smaller CLIPs is too expensive.



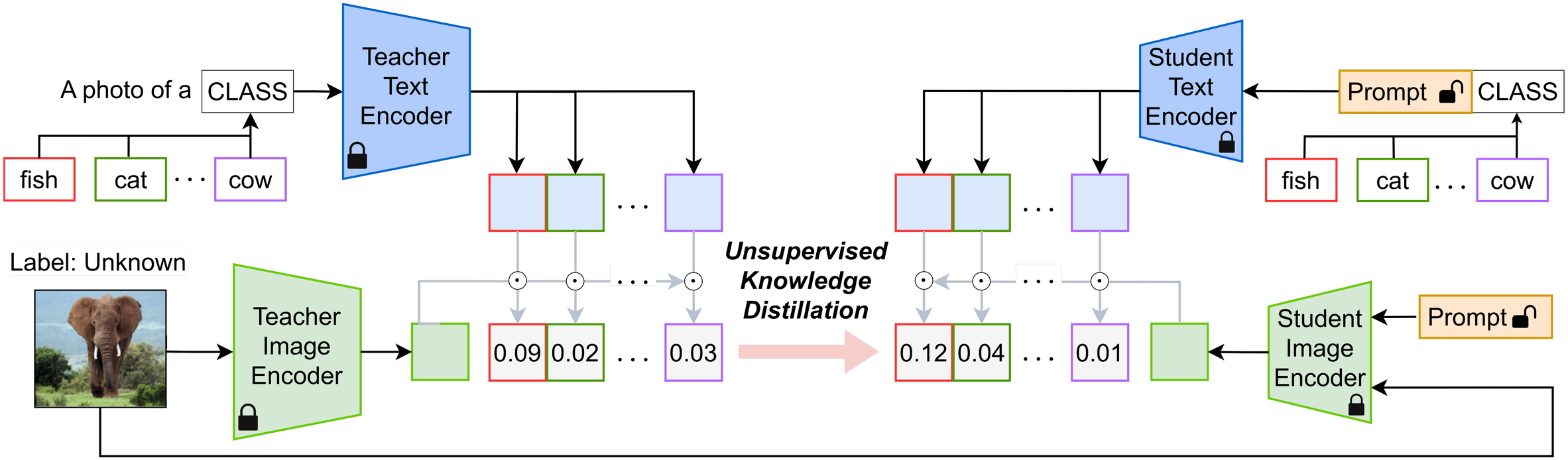
Existing approaches

Prompt Learning is an efficient solution, but the best performing SOTA techniques require ground truth labels.



Knowledge Distillation for Prompt Learning (KDPL)

KDPL can be integrated with any existing prompt learning technique and eliminates the need for labels!



Note that the teacher is frozen and only needed at training time!
Furthermore, distilling at logit level makes KDPL completely architecture independent!

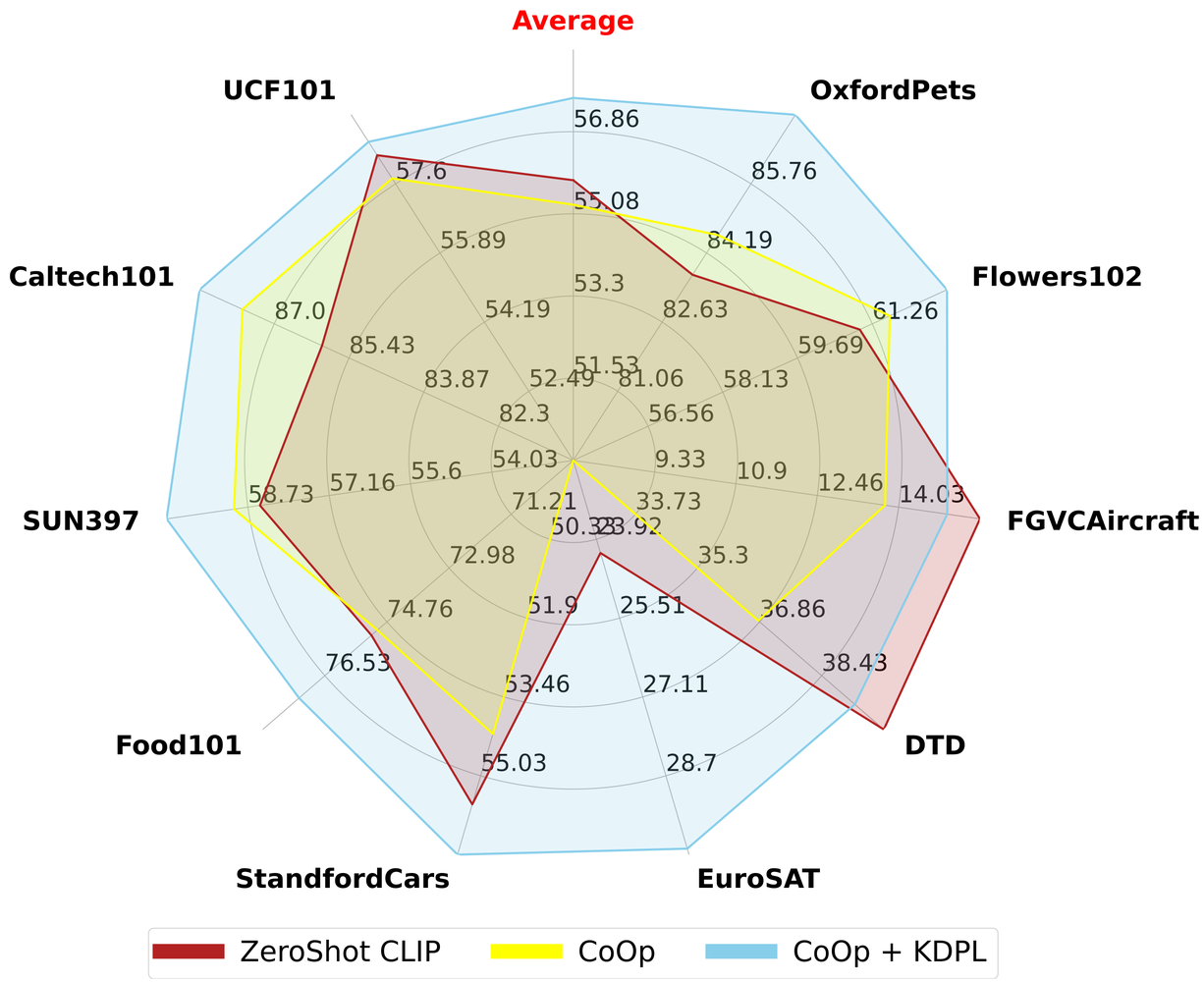
Cross Domain Evaluation

Train 16-shots on ImageNet and evaluate on ImageNet -V2/-Sketch/-A/-R.

Backbone	Method	Source	Target				
		ImageNet	-V2	-S	-A	-R	*Average*
RN50	CLIP (student)	58.20	51.50	33.30	21.70	56.00	40.63
	CoOp	62.40	55.17	33.70	23.13	56.20	42.05
	CoOp + KDPL	**62.73**	**55.37**	**35.20**	**23.27**	**57.77**	**42.90**
	CoCoOp	**63.07**	55.53	34.77	**23.73**	**59.47**	43.38
	CoCoOp + KDPL	62.70	**55.60**	**35.30**	23.43	57.90	**43.06**
ViT-B/32	CLIP (student)	62.00	54.70	40.80	29.60	66.00	47.78
	CoOp	**66.33**	**58.30**	41.40	31.47	65.87	49.26
	CoOp + KDPL	65.97	58.10	**42.50**	**31.63**	**67.37**	**49.90**
	VPT	64.97	56.73	41.27	27.00	66.50	47.88
	VPT + KDPL	**65.10**	**57.37**	**41.67**	**27.77**	**67.47**	**48.57**
	MaPLe	**66.80**	**58.53**	42.23	**30.13**	66.40	49.32
	MaPLe + KDPL	66.50	58.47	**42.77**	29.87	**67.70**	**49.70**
	PromptSRC	**66.33**	**58.70**	42.97	**32.07**	**68.93**	50.67
	PromptSRC + KDPL	66.27	58.60	**43.07**	31.70	68.83	**50.55**
ViT-H/14	CLIP (teacher)	82.80	76.60	71.10	71.10	91.30	77.53

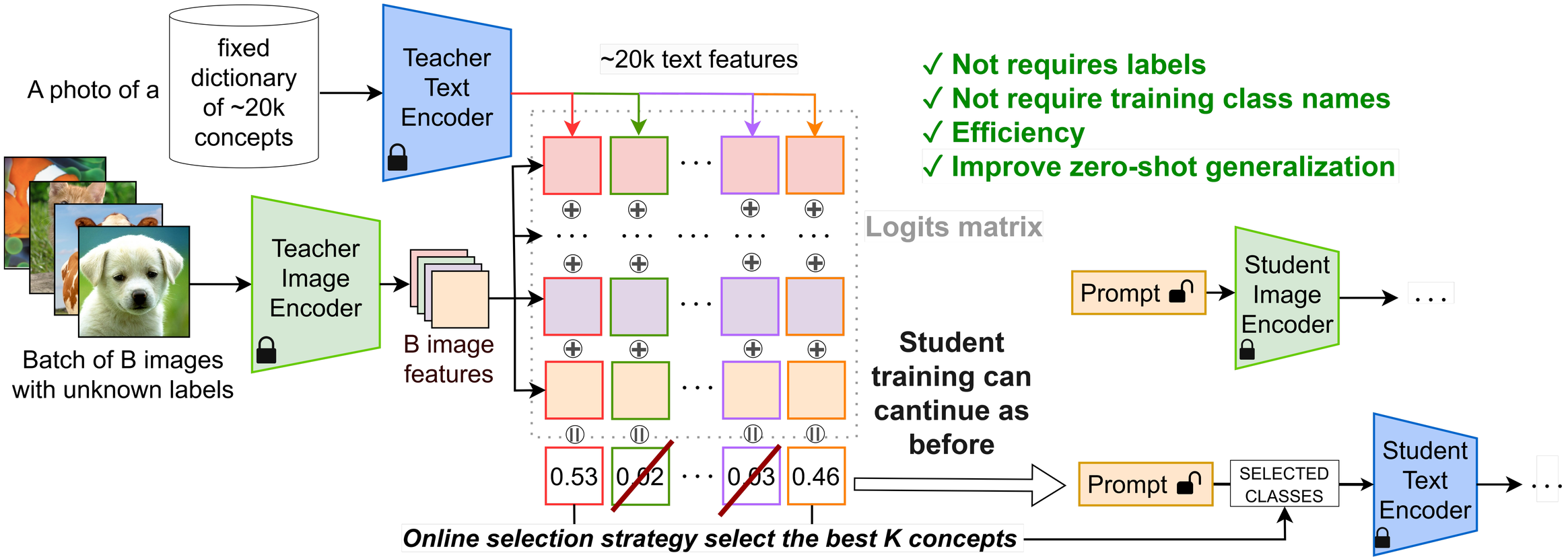
Cross Dataset Evaluation

Train 16 shots on ImageNet and evaluate on 10 different benchmark datasets.



Class-Agnostic Approach: KDPL-CA

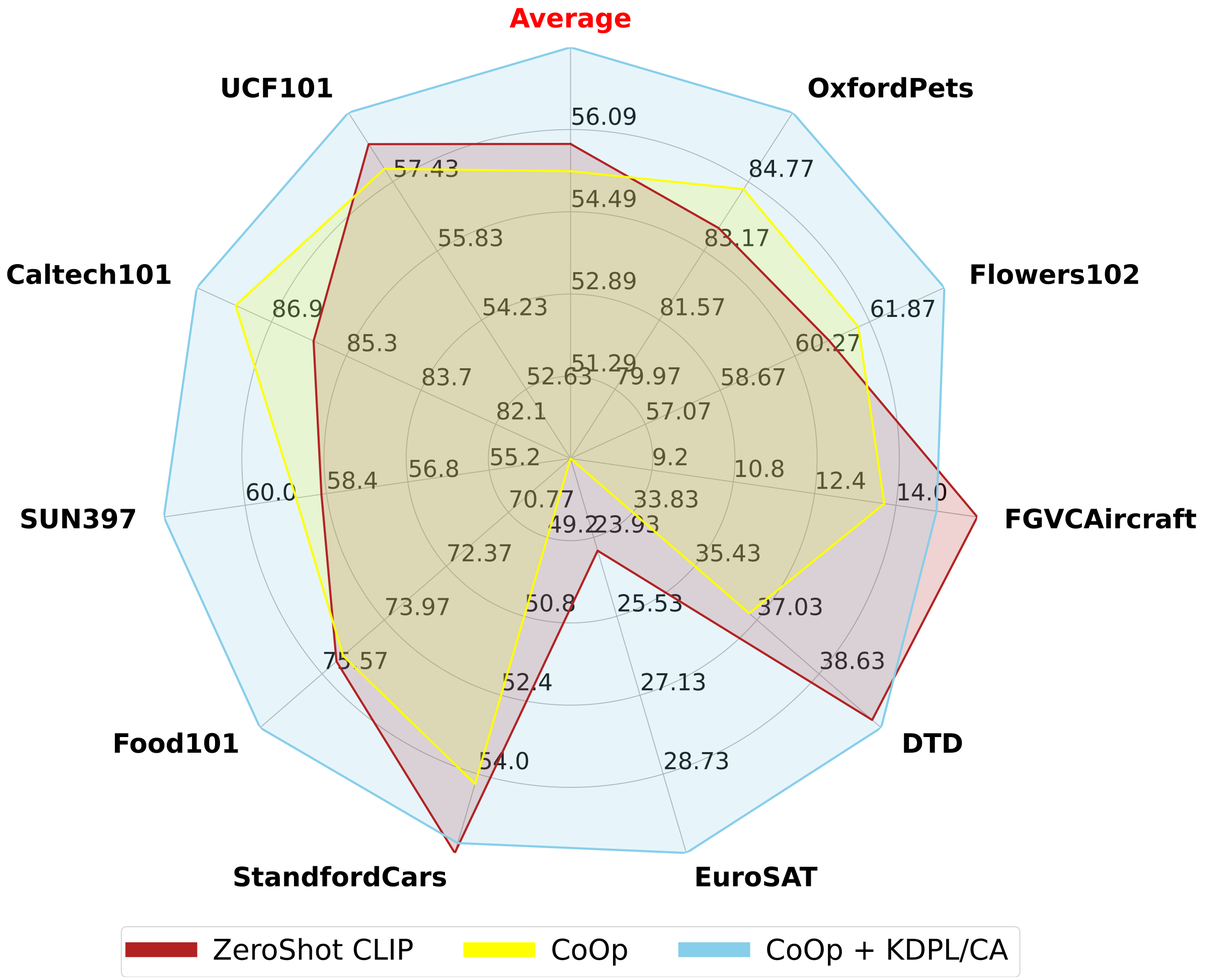
What can we do when we know neither the labels nor the *training class names*?



See section 3.3 of the paper for more details on the Online concepts selection strategy.

KDPL-CA improves zero-shot generalization

Train 16 shots on ImageNet and evaluate on 10 different benchmark datasets.



Improving Zero-shot Generalization of Learned Prompts via Unsupervised Knowledge Distillation

*Marco Mistretta, *Alberto Baldrati, Marco Bertini, Andrew D. Bagdanov

Thank you for your attention!

arXiv

GitHub

MICC
Media Integration and Communication Center

Presented by Andrew D. Bagdanov

ECCV