Table 1: Comparison with existing methods in the base-to-novel generalization setting on 11 datasets with the 16-shot samples from the base classes. HM: Harmonic mean.

	Base	Novel	HM
ProGrad[1] (ICCV'23)	82.48	70.75	76.16
CoPrompt[2] (ICLR'24)	94.00	77.23	80.48
DeKgTCP[3] (ICLR'25)	84.96	76.38	80.44
TAP[4] (ICLR'25)	84.75	77.63	81.04
MMRL[5] (CVPR'25)	85.68	77.16	81.20
CoLD (Ours)	**86.19**	**78.30**	**82.06**

[1] Zhu et al. "*ProGrad: Prompt-aligned Gradient for Prompt Tuning.*"
[2] Roy et al. "*CoPrompt: Consistency-guided prompt learning for vision-language models.*"
[3] Li et al. "*DeKgTCP: Divergence-enhanced knowledge-guided context optimization for visual-language prompt tuning.*"
[4] Ding et al. "*TAP: Tree of Attributes Prompt Learning for Vision-Language Models.*"
[5] Guo et al. "*MMRL: Multi-Modal Representation Learning for Vision-Language Models.*"

Table 2: Comparison of CoLD with existing methods in the domain generalization setting.

	Source	**Target**				
	ImageNet	-V	-S	-A	-S	Avg.
ProGrad (ICCV'23)	**72.24**	64.73	47.61	49.39	74.58	59.08
CoPrompt (ICLR'24)	70.80	64.25	49.43	50.50	77.51	60.42
MMA[1] (CVPR'24)	71.00	64.33	49.13	51.12	77.32	60.48
MMRL (CVPR'25)	72.03	64.47	49.13	**51.20**	77.53	60.58
CoLD (Ours)	71.82	**65.31**	**50.64**	51.07	**77.72**	**61.19**

[1] Yang et al. "*Mma: Multi-modal adapter for vision-language models.*"

Table 3: Comparison of CoLD in few shot classification results with 16 shots.

	16-Shot Classification											
	Average	ImageNet	Caltech	Pets	Cars	Flowers	Food	FGVC	SUN397	DTD	EuroSAT	UCF101
CoOp (IJCV'22)	79.89	71.87	95.57	91.87	83.07	97.07	84.20	43.40	74.67	69.87	84.93	82.23
CoCoOp (CVPR'22)	74.90	70.83	95.16	93.34	71.57	87.84	87.25	31.21	72.15	63.04	73.32	78.14
MaPLe (CVPR'23)	81.79	72.33	96.00	92.83	83.57	97.00	85.33	48.40	75.53	71.33	92.33	85.03
PSRC (CVPR'23)	82.87	73.17	96.07	93.67	83.83	97.60	87.50	50.83	77.23	72.73	92.43	86.47
LLaMP (CVPR'24)	83.81	73.49	97.08	94.21	86.07	98.06	**87.62**	56.07	77.02	74.17	91.31	86.84
TAP (ICLR'25)	83.37	**73.76**	96.73	93.90	85.37	98.10	87.53	50.43	77.30	74.90	91.90	87.17
MMRL (CVPR'25)	84.34	73.40	97.13	93.83	86.43	98.40	87.03	57.60	**77.70**	75.30	**93.37**	87.60
CoLD (Ours)	**84.62**	73.68	**97.17**	**94.65**	**87.39**	**98.72**	87.58	**58.79**	77.11	**76.56**	91.29	**87.91**

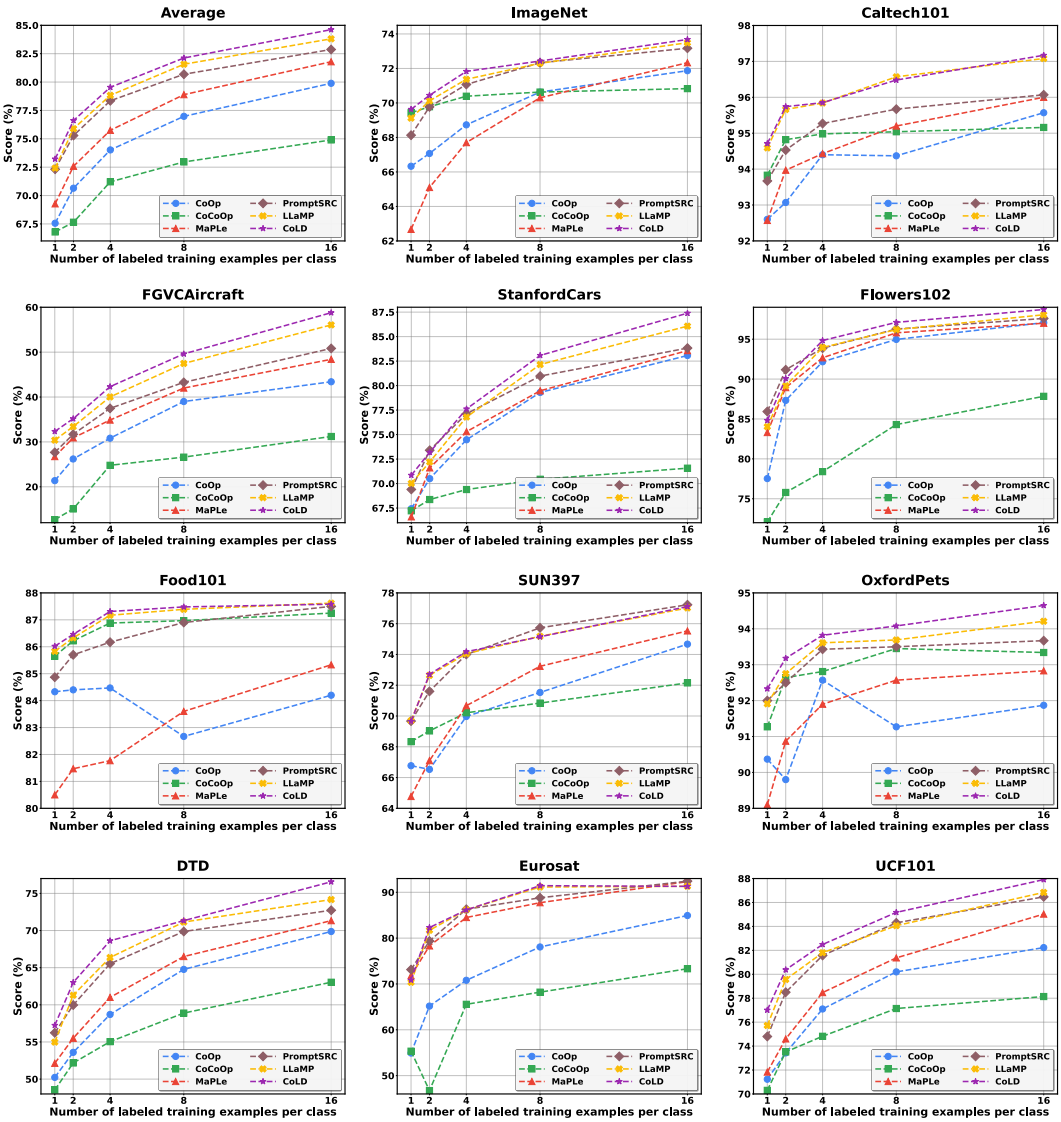

Figure 1: Comparison of CoLD with previous methods on few-shot learning across 11 datasets.

Table 4: Comparison in the cross-dataset evaluation. The model is trained on the entire class of ImageNet (16 shots) and evaluated on the other 10 datasets.

	Source	**Target**										
	ImageNet	Caltech	Pets	Cars	Flowers	Food	FGVC	SUN397	DTD	EuroSAT	UCF101	Average
CoOp	71.51	93.70	89.14	64.51	68.71	85.30	18.47	64.15	41.92	46.39	66.55	63.88
MaPLe	70.72	93.53	90.49	65.57	72.23	86.20	24.74	67.01	46.49	48.06	68.69	66.30
PSRC	71.27	93.60	90.25	65.70	70.25	86.15	23.90	67.10	46.87	45.50	68.75	65.81
ProGrad	72.24	91.52	89.64	62.39	67.87	85.40	20.16	62.47	36.42	43.46	64.29	62.36
CoPrompt	70.80	94.50	90.73	65.67	72.30	86.43	24.00	67.57	47.07	51.90	**69.73**	67.00
DeKgTCP	**72.33**	94.73	90.02	65.49	72.39	**86.59**	25.05	67.19	44.47	51.37	68.78	66.61
TAP	72.30	94.30	90.70	65.60	70.93	86.10	24.57	**68.30**	**50.20**	46.00	68.90	66.56
MMRL	72.03	94.67	**91.43**	66.10	72.77	86.40	26.30	67.57	45.90	53.10	68.27	67.25
CoLD	71.82	**94.75**	91.13	**66.25**	**72.86**	86.21	**26.51**	67.64	47.23	**53.14**	68.96	**67.47**