

# ID-like Prompt Learning for Few-Shot Out-of-Distribution Detection

## Supplementary Material

### 1. Appendix

#### 1.1. More experiments in different scenarios.

We add more experiments to demonstrate the broad application scope of the proposed method. Experimental results demonstrate its excellent performance on wide scenarios.

Table 1. OOD detection performance for different datasets as ID.

Method	OOD Dataset					
	iNaturalist		Places		Texture	
	FPR95↓	AUROC↑	FPR95↓	AUROC↑	FPR95↓	AUROC↑
ID: CUB-200						
MCM	9.83	98.24	6.65	98.57	6.97	98.75
CoOp	13.07	97.80	4.32	98.99	2.30	99.55
Ours	0.04	99.96	0.50	99.85	0.00	99.98
ID: Oxford-Pet						
MCM	2.85	99.38	2.11	99.56	0.80	99.81
CoOp	5.11	99.00	1.43	99.73	0.53	99.85
Ours	0.00	99.97	0.27	99.89	0.04	99.98
ID: Food-101						
MCM	0.64	99.78	1.86	99.58	4.04	98.62
CoOp	2.45	99.49	1.29	99.70	4.41	98.34
Ours	0.28	99.80	0.08	99.72	0.28	99.89
ID: Stanford-Cars						
MCM	0.05	99.77	0.24	99.89	0.02	99.96
CoOp	0.10	99.79	0.20	99.91	0.00	99.97
Ours	0.03	99.60	0.22	99.62	0.07	99.70

#### 1.2. Experiments on near/hard OOD.

For validating near OOD detection, we employ the NINCO dataset [?] proposed in recent work. This dataset is specifically designed to create a cleaner OOD benchmark for ImageNet, making it more challenging to differentiate from ImageNet compared to iNaturalist and Texture datasets, as shown in Table 2. The proposed method still achieves significant improvements.

Table 2. OOD detection performance for ImageNet-1k as ID, NINCO [?] as OOD.

OOD Dataset	Method					
	MCM		CoOp		Ours	
	FPR95↓	AUROC↑	FPR95↓	AUROC↑	FPR95↓	AUROC↑
NINCO [?]	79.28	74.36	82.17	71.64	71.73	77.69

#### 1.3. The effectiveness of ID-like prompts.

We compare the average performance of our method in OOD detection on ImageNet-1k under different conditions. We show the MCM scores when only ID prompts are presented (ID prompts), MCM scores when both ID prompts and OOD prompts are used simultaneously (All prompts), and the

Table 3. Ablation study of the effectiveness of prompts and the proposed OOD score function.

Method	Average			
	One-shot		Four-shot	
	FPR95↓	AUROC↑	FPR95↓	AUROC↑
ID prompts	49.57	88.25	49.28	87.86
All prompts	44.18	89.80	42.51	89.81
Ours	27.69	94.29	26.08	94.36

proposed score when all prompts are provided (Ours). As shown in Table 3, we observe that with the introduction of OOD prompts, the OOD detection performance significantly improves. This indicates that OOD prompts can promisingly distinguish between ID and OOD samples.