

# Supplementary Materials: Rethinking the Effect of Uninformative Class Name in Prompt Learning

Fengmao Lv  
Southwest Jiaotong University  
Chengdu, China  
fengmaolv@swjtu.edu.cn

Changru Nie  
Southwest Jiaotong University  
Chengdu, China  
niechangru@my.swjtu.edu.cn

Jiayang Zhang\*  
University of Electronic Science and  
Technology of China  
Chengdu, China  
jiayangzhang@std.uestc.edu.cn

Guowu Yang  
University of Electronic Science and  
Technology of China  
Chengdu, China  
guowu@uestc.edu.cn

Guosheng Lin  
Nanyang Technological University  
Singapore, Singapore  
gslin@ntu.edu.sg

Xiao Wu  
Southwest Jiaotong University  
Chengdu, China  
wuxiaohk@swjtu.edu.cn

Tianrui Li  
Southwest Jiaotong University  
Chengdu, China  
trli@swjtu.edu.cn

## A Additional Result Comparisons

In order to further verify the applicability and the effectiveness of IDAPL, we conduct additional comparison studies with the very recent prompt learning approach CoPrompt [3]. CoPrompt significantly improves the generalization ability of prompt learning paradigm by learning the consistency of the language and image branches with the GPT-generated class descriptions and the augmented images, respectively. Notably, for fair comparisons, we set  $\alpha = 0.1$  &  $\beta = 5$  as the same as implementing IDAPL under the CoOp and MaPLe baselines. The rest hyper-parameter settings follow the original implementation of CoPrompt. In addition, the results of CoPrompt are implemented based on its official open source code. The results are shown in Table A1. We can see that IDAPL achieves performance improvements on 7 of 11 datasets on both the base and novel classes, with improvements on 8 of 11 datasets on the harmonic mean of accuracies on base and novel classes. These results further verify the effectiveness of IDAPL and the broad applicability of IDAPL on existing prompt learning approaches.

## B Comparing with GPT enhanced baselines

To further validate the improvements of our approach, we conduct additional comparison studies with GPT enhanced baseline models under the base-to-novel setting. Specifically, we combine GPT generated class descriptions collected by Maniparambil et al. [2] with baseline model CoOp [4] and MaPLe [1] following our implementation introduced in Section 3.2. The comparison results are shown in Table A2. It is clear that, in general, our approach can achieve better performances in both base and novel classes. Specifically, comparing with directly combing GPT descriptions with CoOp, applying our IDAPL approach into the training paradigm of CoOp achieves better performances on 8 out of 11 datasets for base classes and 9 out of 11 datasets for novel classes. Moreover, comparing with combing GPT descriptions with MaPLe, applying our IDAPL

**Table A1: Comparison with CoPrompt on the base-to-novel generalization setting. H indicates the harmonic mean of accuracies on base and novel categories.**

	CoPrompt			CoPrompt+IDAPL		
	Base	Novel	H	Base	Novel	H
Average	83.0	74.7	78.6	<b>84.2</b>	<b>76.3</b>	<b>80.1</b>
ImageNet	76.5	<b>71.3</b>	73.8	<b>77.0</b>	71.2	<b>74.0</b>
FGVCAircraft	35.7	33.8	34.7	<b>44.0</b>	<b>36.7</b>	<b>40.0</b>
Food101	<b>90.5</b>	91.5	<b>91.0</b>	89.7	<b>91.8</b>	90.7
Caltech101	98.6	<b>95.4</b>	<b>97.0</b>	<b>99.1</b>	94.5	96.8
Flowers102	96.9	74.6	84.3	<b>97.1</b>	<b>74.8</b>	<b>84.5</b>
OxfordPets	<b>96.4</b>	<b>97.7</b>	<b>97.0</b>	95.0	96.9	95.9
StanfordCars	72.6	69.3	70.9	<b>78.4</b>	<b>73.1</b>	<b>75.6</b>
SUN397	82.5	<b>79.8</b>	81.1	<b>83.5</b>	<b>79.8</b>	<b>81.6</b>
DTD	<b>83.9</b>	58.5	68.9	83.0	<b>62.2</b>	<b>71.1</b>
EuroSAT	93.5	71.3	80.9	<b>94.2</b>	<b>79.2</b>	<b>86.0</b>
UCF101	<b>85.9</b>	78.4	82.0	<b>85.9</b>	<b>78.6</b>	<b>82.1</b>

approach into MaPLe achieves better performances on 8 out of 11 datasets and equivalent performances on 2 of 11 datasets for base classes, as well as better performances on 7 out of 11 datasets and equivalent performances on 1 out of 11 datasets for novel classes. These performance benefits clearly verify the effectiveness of our approach, and that the improvement achieved by IDAPL doesn't solely come from the usage of GPT.

\*Corresponding author.

**Table A2: Comparison with GPT enhanced baseline models on the base-to-novel generalization setting. H indicates the harmonic mean of accuracies on base and novel categories. The blue texts indicate the performance differences between our approach and directly combing GPT descriptions with baseline models.**

(a) Average over 11 datasets				(b) ImageNet				(c) FGVCaircraft			
	Base	Novel	H		Base	Novel	H		Base	Novel	H
CoOp	82.7	63.2	71.7	CoOp	76.5	67.9	71.9	CoOp	40.4	22.3	28.7
CoOp+GPT	82.5	72.1	76.9	CoOp+GPT	76.3	69.6	72.8	CoOp+GPT	37.6	32.5	34.9
CoOp+IDAPL	84.1	74.6	79.1	CoOp+IDAPL	76.0	69.7	72.7	CoOp+IDAPL	44.3	29.1	35.1
	+1.6	+2.5	+2.1		-0.3	0.1	-0.1		+6.7	-3.4	+0.3
MaPLe	82.3	75.1	78.5	MaPLe	76.7	70.5	73.5	MaPLe	37.4	35.6	36.5
MaPLe+GPT	82.2	76.6	79.3	MaPLe+GPT	76.3	71.0	73.6	MaPLe+GPT	37.8	37.1	37.4
MaPLe+IDAPL	84.2	77.7	80.8	MaPLe+IDAPL	76.8	70.8	73.7	MaPLe+IDAPL	44.0	35.9	39.6
	+2.0	+1.1	+1.5		+0.5	-0.3	+0.1		+6.2	-1.2	+2.1
(d) Food101				(e) Caltech101				(f) Flowers102			
	Base	Novel	H		Base	Novel	H		Base	Novel	H
CoOp	88.3	82.3	85.2	CoOp	98.0	89.8	93.7	CoOp	97.6	59.7	74.1
CoOp+GPT	90.6	90.6	90.6	CoOp+GPT	98.0	95.9	96.9	CoOp+GPT	96.4	72.8	83.0
CoOp+IDAPL	89.9	91.9	90.9	CoOp+IDAPL	98.5	95.2	96.8	CoOp+IDAPL	98.1	77.1	86.3
	-0.7	+1.3	+0.3		+0.5	-0.7	-0.1		+1.7	+4.3	+3.4
MaPLe	90.7	92.1	91.4	MaPLe	97.7	94.4	96.0	MaPLe	95.9	72.5	82.6
MaPLe+GPT	90.9	91.9	91.4	MaPLe+GPT	98.5	95.9	97.1	MaPLe+GPT	95.3	74.5	83.7
MaPLe+IDAPL	90.9	92.1	91.5	MaPLe+IDAPL	97.8	94.2	96.0	MaPLe+IDAPL	96.7	76.4	85.3
	0.0	+0.2	+0.1		-0.7	-1.6	-1.2		+1.4	+1.8	+1.7
(g) OxfordPets				(h) StanfordCars				(i) SUN397			
	Base	Novel	H		Base	Novel	H		Base	Novel	H
CoOp	93.7	95.3	94.5	CoOp	78.1	60.4	68.1	CoOp	80.6	65.9	72.5
CoOp+GPT	95.6	97.0	96.3	CoOp+GPT	76.7	74.4	75.5	CoOp+GPT	81.3	74.0	77.5
CoOp+IDAPL	94.8	97.3	96.0	CoOp+IDAPL	81.7	75.9	78.7	CoOp+IDAPL	81.9	79.5	80.7
	-0.8	+0.3	-0.3		+5.0	+1.5	+3.2		+0.6	+5.5	+3.2
MaPLe	95.4	97.8	96.6	MaPLe	72.9	74.0	73.5	MaPLe	80.8	78.7	79.7
MaPLe+GPT	95.5	97.7	96.6	MaPLe+GPT	71.4	73.5	72.4	MaPLe+GPT	80.7	78.9	79.8
MaPLe+IDAPL	95.5	97.7	96.6	MaPLe+IDAPL	80.8	75.5	78.1	MaPLe+IDAPL	81.9	79.5	80.7
	0.0	0.0	0.0		+9.4	+2.1	+5.6		+1.2	+0.6	+0.9
(j) DTD				(k) EuroSAT				(l) UCF101			
	Base	Novel	H		Base	Novel	H		Base	Novel	H
CoOp	79.4	41.2	54.2	CoOp	92.2	54.7	68.7	CoOp	84.7	56.1	67.5
CoOp+GPT	82.6	51.9	63.7	CoOp+GPT	87.5	64.5	74.3	CoOp+GPT	84.6	70.1	76.7
CoOp+IDAPL	83.8	62.0	71.3	CoOp+IDAPL	90.9	64.7	75.6	CoOp+IDAPL	85.3	78.7	81.9
	+1.2	+10.1	+7.5		+3.4	+0.2	+1.3		+0.7	+8.6	+5.2
MaPLe	80.4	59.2	68.2	MaPLe	94.1	73.2	82.4	MaPLe	83.0	78.7	80.8
MaPLe+GPT	81.6	60.6	69.6	MaPLe+GPT	93.4	81.4	87.0	MaPLe+GPT	83.0	79.7	81.3
MaPLe+IDAPL	83.4	65.3	73.3	MaPLe+IDAPL	93.9	85.8	89.7	MaPLe+IDAPL	84.4	81.4	82.9
	+1.8	+4.7	+3.7		+0.5	+4.4	+2.7		+1.4	+1.7	+1.6

## References

- [1] Muhammad Uzair Khattak, Hanoona Abdul Rasheed, Muhammad Maaz, Salman H. Khan, and Fahad Shahbaz Khan. 2023. MaPLe: Multi-modal Prompt Learning. In *CVPR*. 19113–19122.
- [2] Mayug Maniparambil, Chris Vorster, Derek Molloy, Noel Murphy, Kevin McGuinness, and Noel E. O’Connor. 2023. Enhancing CLIP with GPT-4: Harnessing Visual Descriptions as Prompts. In *ICCV*. 262–271.
- [3] Shuvendu Roy and Ali Etemad. 2023. Consistency-guided Prompt Learning for Vision-Language Models. *CoRR* abs/2306.01195 (2023). <https://doi.org/10.48550/ARXIV.2306.01195> arXiv:2306.01195
- [4] Kaiyang Zhou, Jingkang Yang, Chen Change Loy, and Ziwei Liu. 2022. Conditional Prompt Learning for Vision-Language Models. In *CVPR*. 16795–16804.