Read-only Prompt Optimization for Vision-Language Few-shot Learning (ICCV 2023)

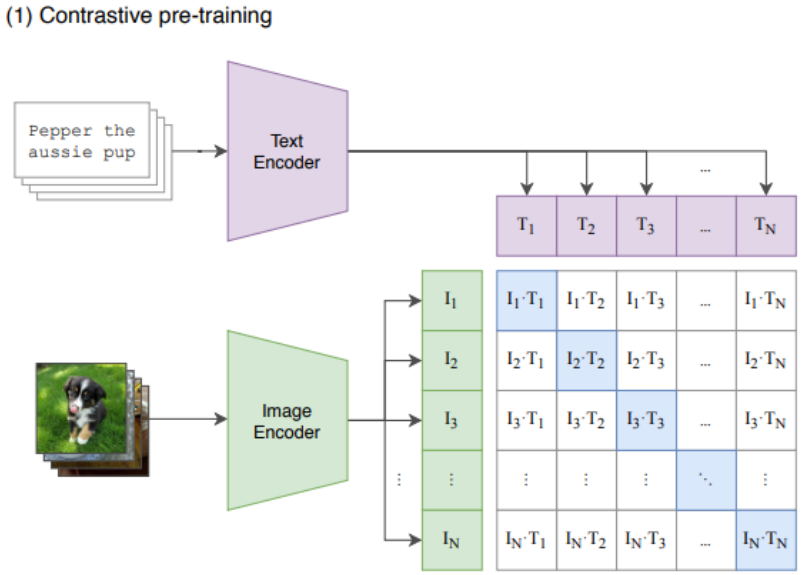
Dongjun Lee∗ Seokwon Song∗ Jihee Suh
Joonmyeong Choi Sanghyeok Lee Hyunwoo J. Kim†
Korea University

Yurim Lee

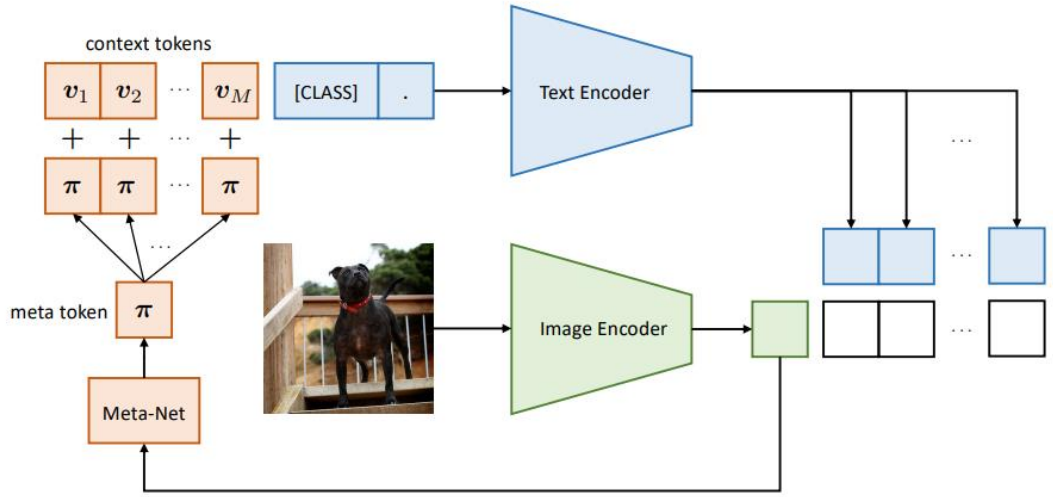
yurimmy65@gmail.com

2024. 2. 29

I . Recap

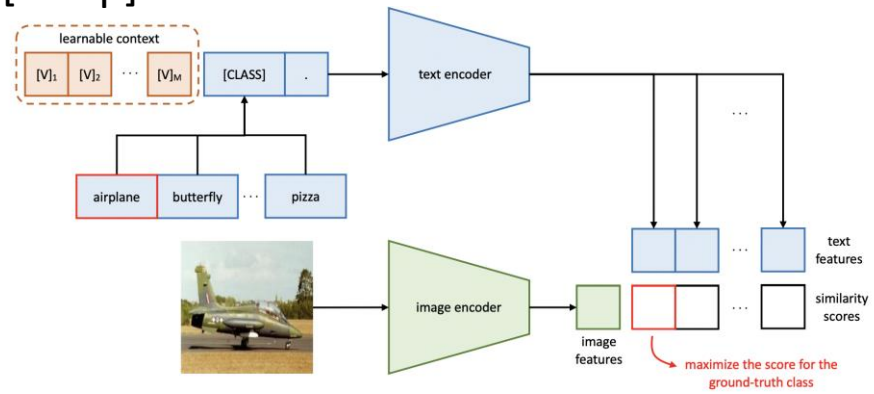


[CLIP]



[CoCoOp]

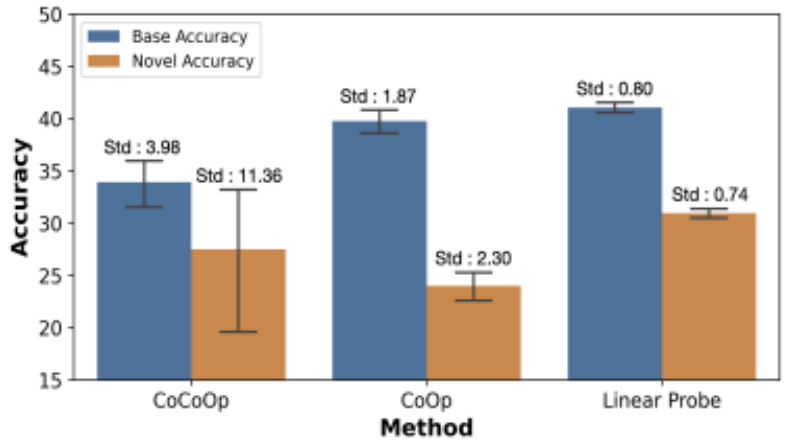
[CoOp]



I . Recap

(a) Conventional Prompt Tuning (b) Linear Probing (c) Read-only Prompt Optimization

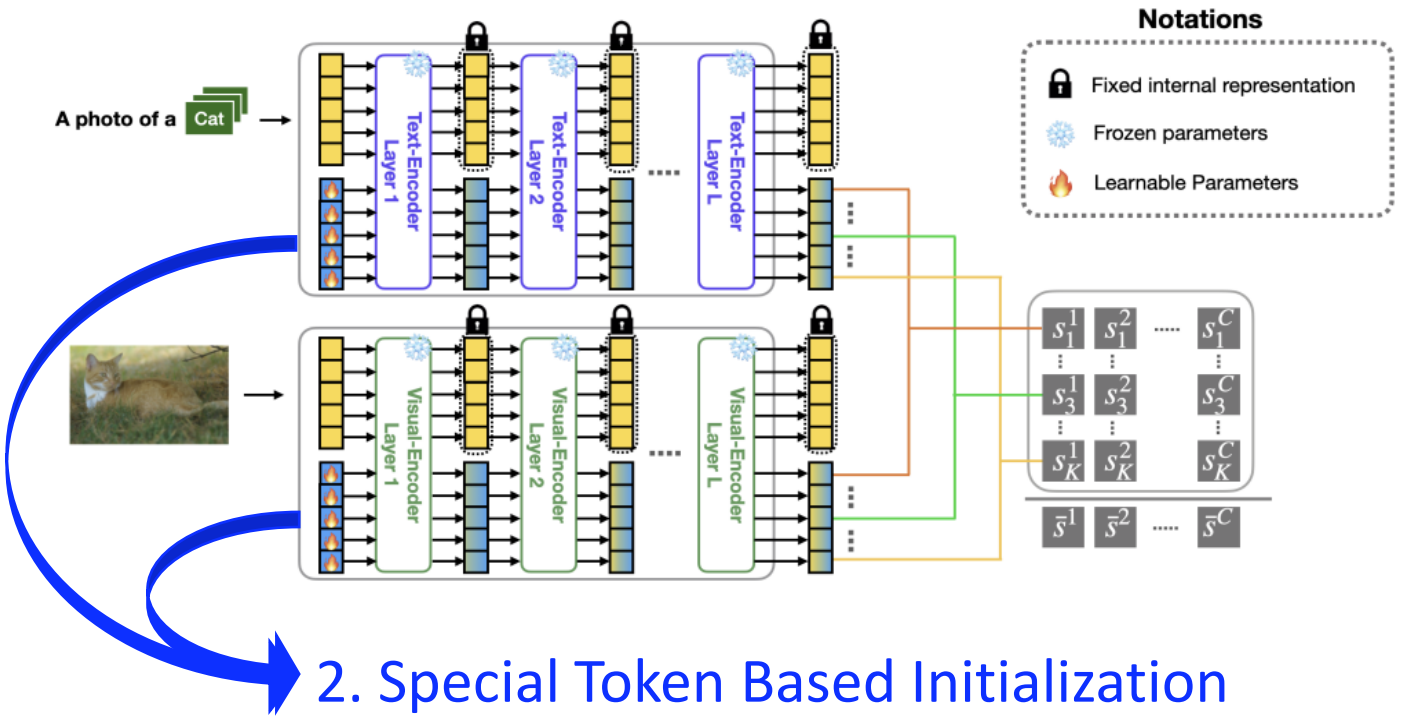
"Internal Representation Shift"



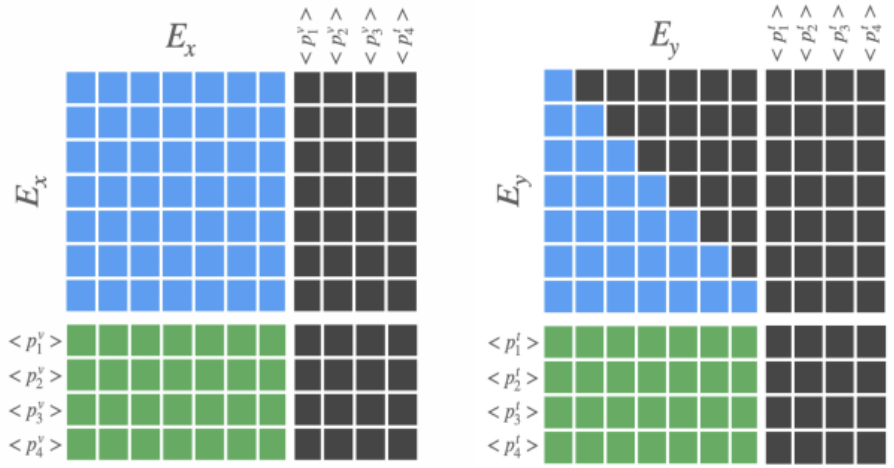
CoOp, CoCoOp: High Variance
-> may have negatively impact robustness & generalization in data-deficient setting

Linear probing: Parameter inefficient(262k), Lack of generalizability in domain-shift task

I . Recap



2. Special Token Based Initialization



(a) Visual Attention Mask (b) Textual Attention Mask

Figure 4: The visualization of attention masks for each encoder.

1. Masked Attention

-> NO internal representation shift

Ⅱ. Experiments

1. K (Num of Prompts) 개수에 따른 성능 변화

2. Epoch에 따른 성능 변화

3. Initial Prompting에 따른 성능 변화

4. Seed에 따른 variance 변화

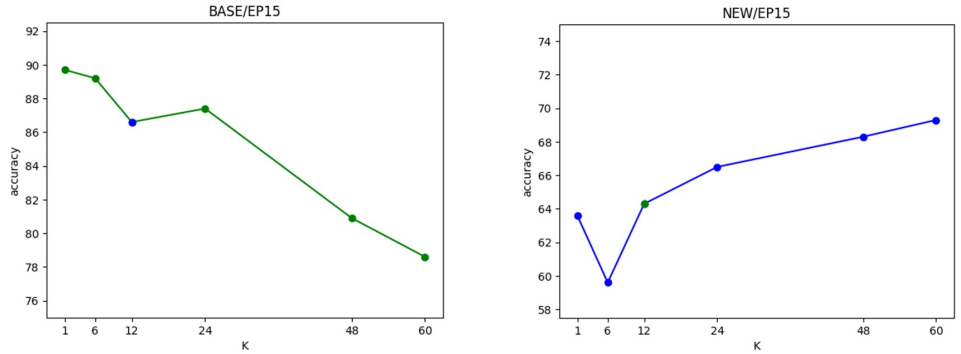
Ⅱ. Experiments

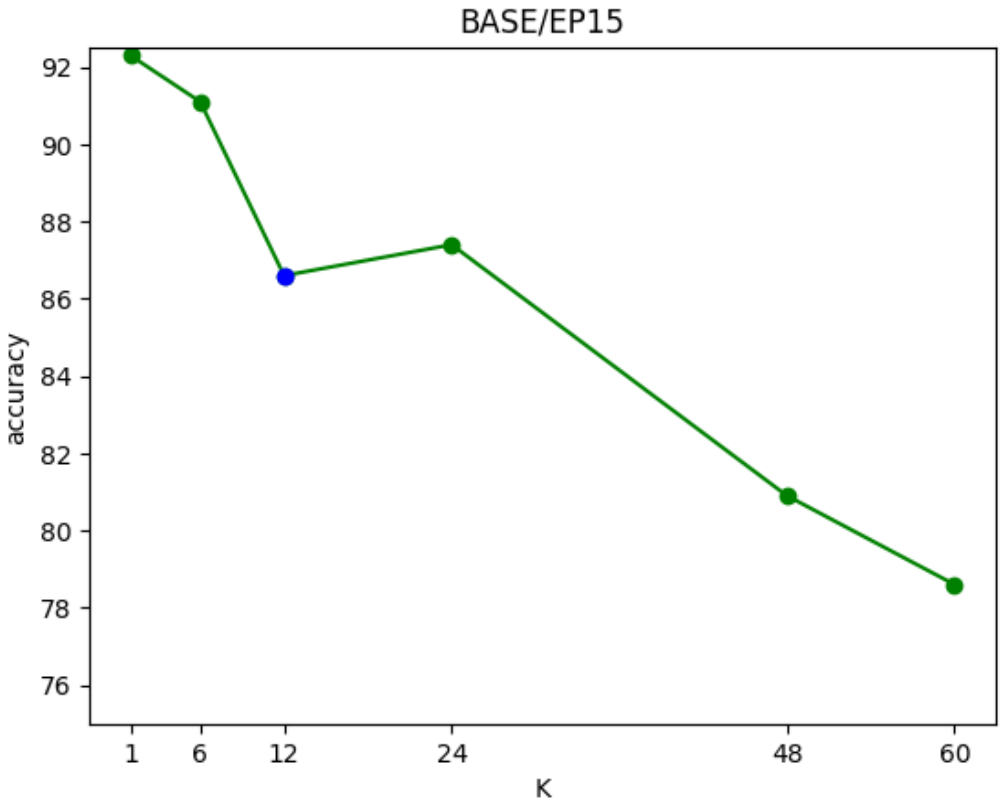
Eurosat[10classes, 27000 images] **(sealake, highway, forest..)**

X16 shots

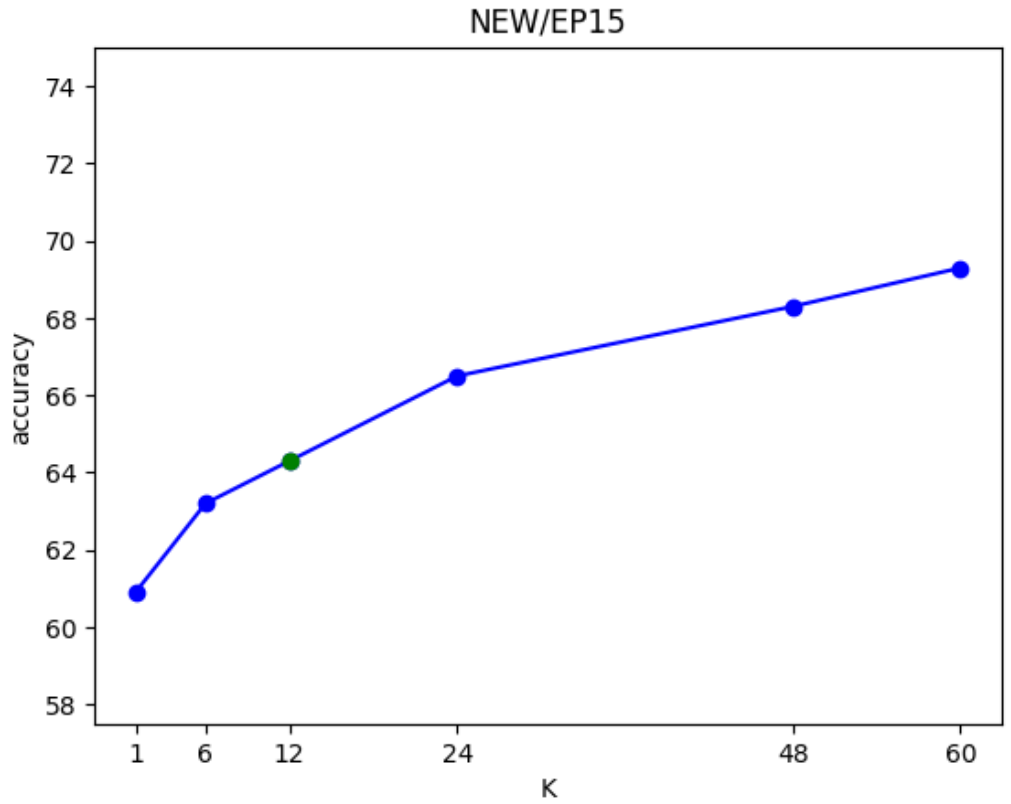
Ⅱ. Experiments

2.1 K(Num of Prompts) 개수에 따른 성능 변화





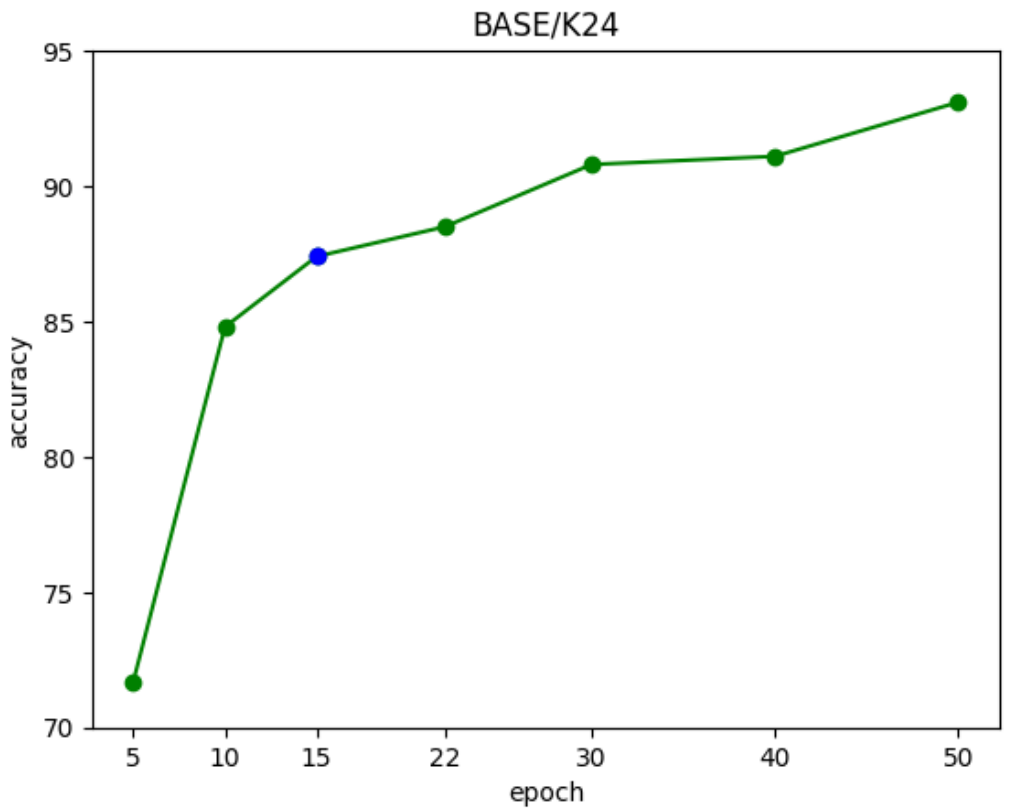
Too many cooks spoil the broth



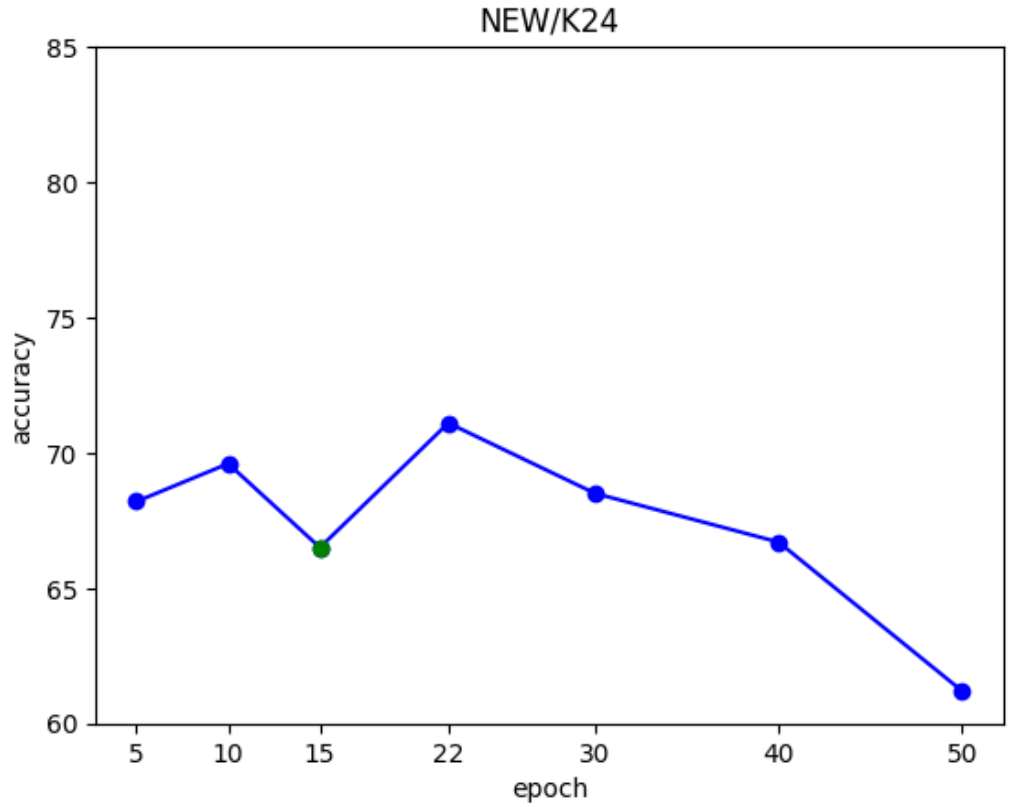
Two heads are better than one

Ⅱ. Experiments

2.2 Epoch 수에 따른 성능 변화



The more you learn,
the more intelligent you get



Learning more doesn’t always mean
you get more intelligent

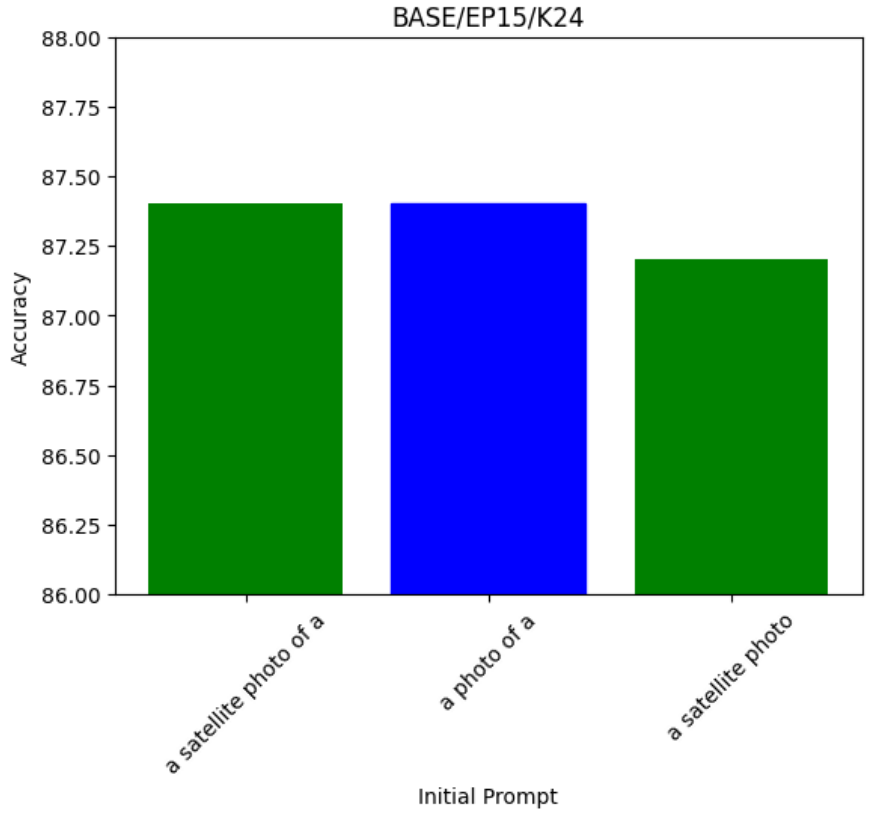
Ⅱ. Experiments

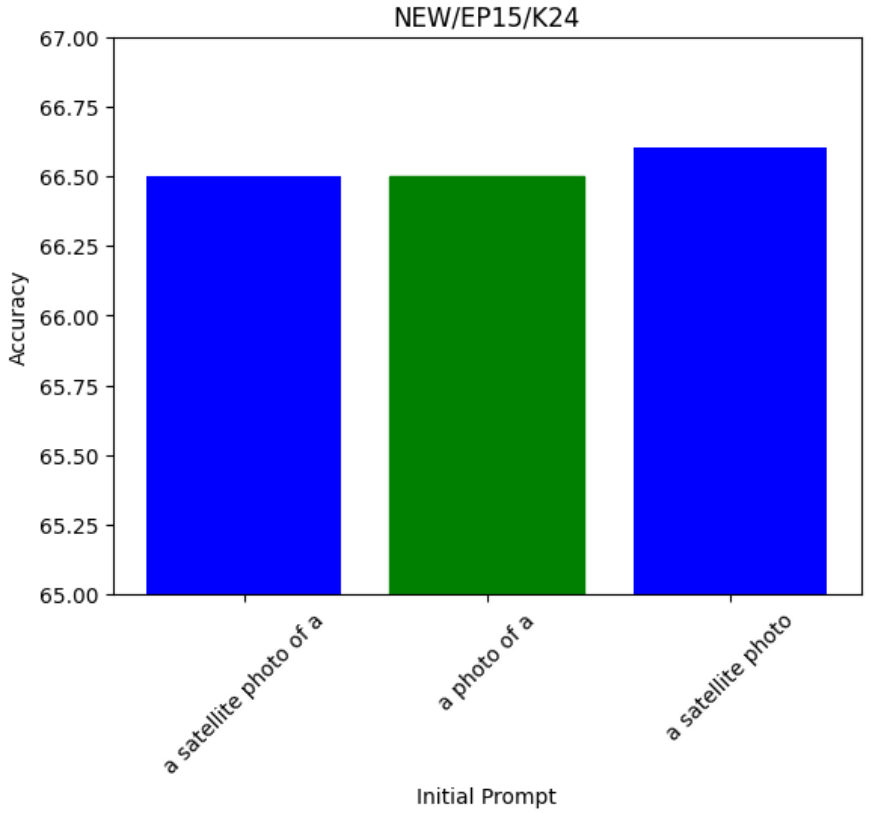
2.3 Initial Prompt에 따른 성능 변화

EuroSAT	Prompt	Accuracy
	a photo of a [CLASS].	24.17
	a satellite photo of [CLASS].	37.46
	a centered satellite photo of [CLASS].	37.56
	$[V]_1$ $[V]_2$ ... $[V]_M$ [CLASS].	**83.53**

(d)

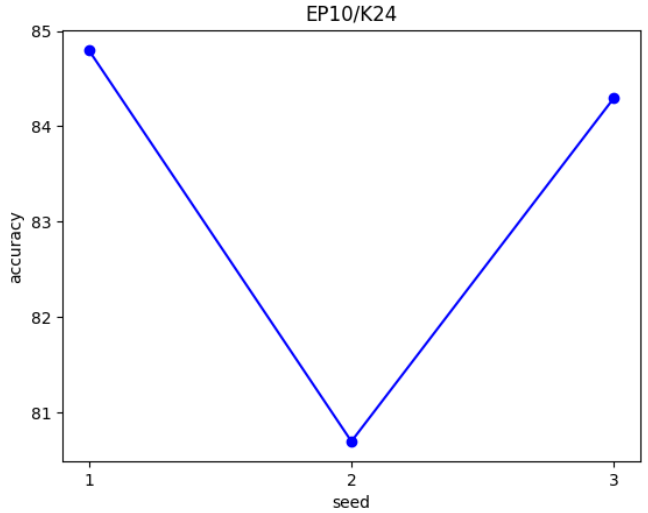
[CoOp]

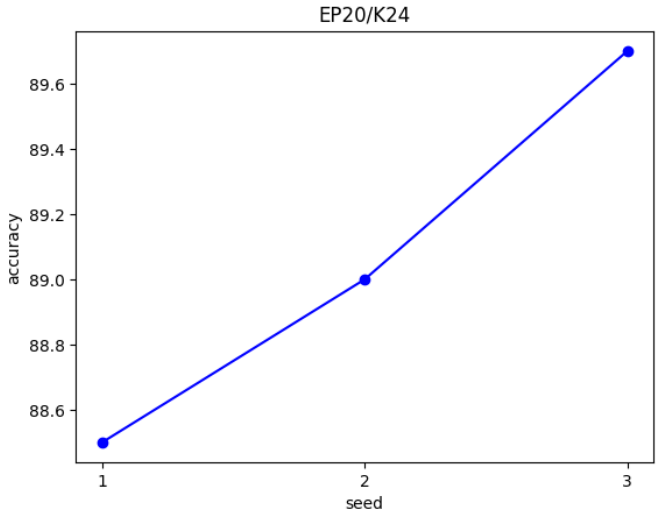


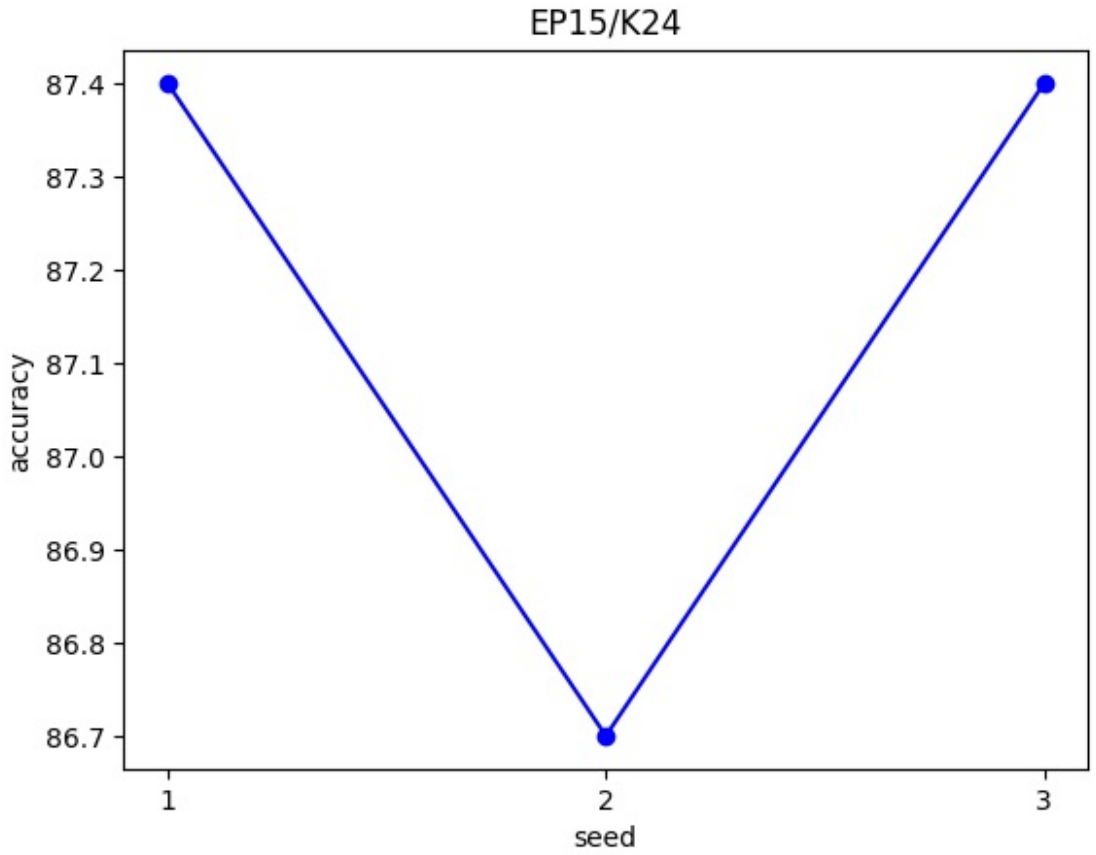


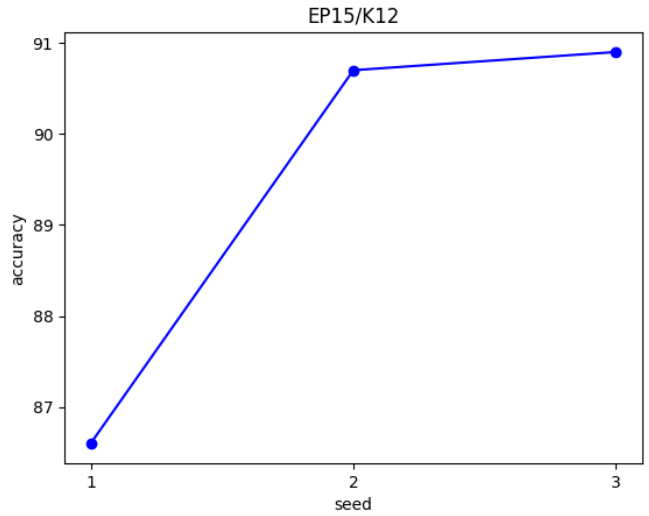
Ⅱ. Experiments

2.4 seed에 따른 variance 변화









Ⅲ. Limitations

1. GPU와 Time Limit으로 인해 seed1로 진행
   -> seed 다양화를 통한 variance issue 극복 x

2. Base to New Generalization만 실험
   -> Maple과 같이 Cross Dataset Evaluation을 통한 'Generalizability'에 관한 성능 평가 다양화의 필요

3. 새로운 Idea의 적용보다는, Analysis에 그침

Ⅳ. Hope Works

1. CoOp과 같이 class별로 Prompt를 달리 하는 specific class와 모든 class에 동일한 prompt를 부여하는 unified class task 진행

2. Image, Text Prompt 다르게 제공 -> 어떤 방식으로 다르게 주어야 할까? 에 관한 고민 필요

Thank you!

https://arxiv.org/pdf/2308.14960v1.pdf
https://github.com/mlvlab/RPO