Taha Koleilat* Hojat Asgariandehkordi Hassan Rivaz Yiming Xiao

{taha.koleilat, hojat.asgariandehkordi, hassan.rivaz, yiming.xiao}@concordia.ca

????????????VLM????
CLIP?????????????????????????????????? VLM
??
??
??????????CoOp??????????????????????????????
BiomedCoOp????????????? BiomedCLIP
??
???con-???????????????????????????? (LLM)
?????????????????????????????????????? 9
???? 11 ?????????? 10
??
??https://github.com/HealthX-Lab/BiomedCoOp?

1. Introduction

The latest breakthroughs in vision-language models (VLMs) have opened new possibilities for leveraging multi-modal data in diverse applications. Unlike traditional supervised learning that focuses on closed-set visual concepts, models like Contrastive Language-Image Pre-training (CLIP) [37], which align visual and textual information through contrastive pre-training, allow the exploration of open-set visual concepts, thanks to the adoption of natural language supervision. However, the success of these models often relies heavily on the quality of the textual prompts that guide their predictions while full-model fine-tuning for large-scale VLMs is impractical. To mitigate these, prompt learning that optimizes textual prompts in vision-language models [25, 50, 51] has emerged as one of the critical techniques to enhance performance without the need for extensive fine-tuning. Notably, the pioneering work of Context Optimization (CoOp) [51] introduced this approach for CLIP by treating text prompts as learnable context vectors and preserving the pre-trained model weights. Meanwhile, other approaches [16, 19, 47] focus on lightweight few-shot adaptation through Adapters [18] and Linear Probes [37] to offer parameter-efficient solutions for model adaptation in downstream tasks.

Different from natural images, biomedical images include a wide range of contrasts and modalities, depending on the image acquisition devices and parameters. These images, such as MRI and ultrasound, often have unique visual appearances that can be more difficult to interpret than typical photographs. In addition, image features (e.g., color, texture, shape, and anatomical context) that are related to physiological and pathological changes are more nuanced and complex to describe, and can differ between image modalities. Finally, due to privacy concerns and the high requirement for clinical expertise, large datasets of well-annotated biomedical images are scarce for developing clinical deep learning models. While VLMs and the associated prompt learning techniques have shown success across natural image datasets and benchmarks, their application in the biomedical imaging domain (e.g., diagnosis), which has distinct challenges, remains largely under-explored.

Due to the unique domain knowledge of biomedical images, the backbone vision-language model for prompt learning may require tailored pre-training for the best outcome. Biomed-specific VLMs, such as BiomedCLIP [48]—pretrained on 15 million biomedical image-text pairs from internet resources—are better suited for biomedical tasks

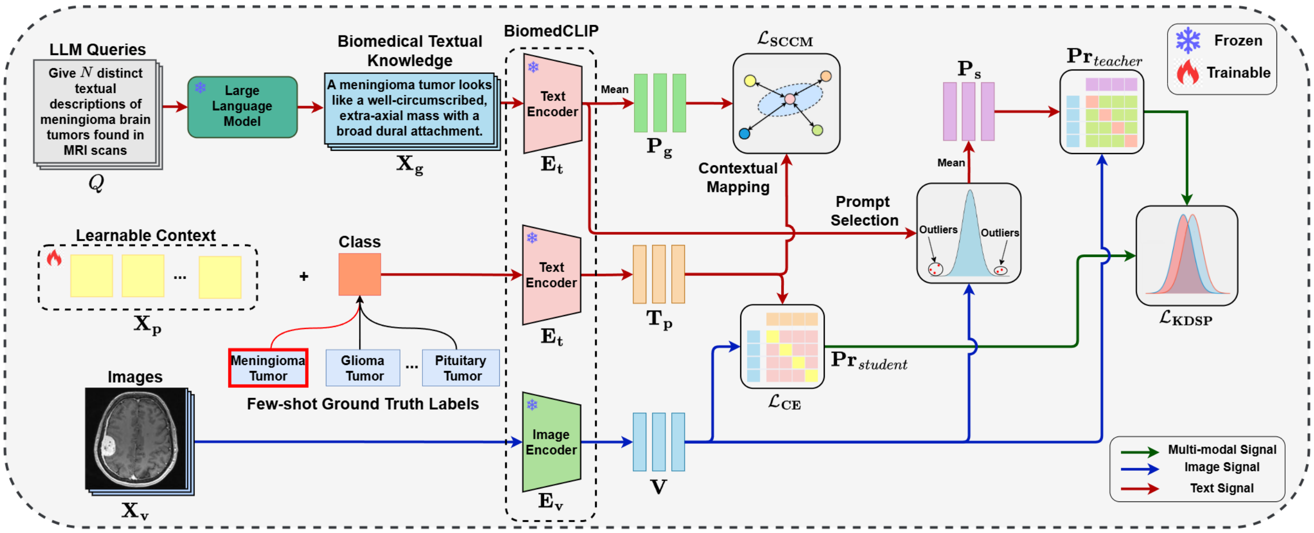
*Corresponding author

and robustness across a wide range of medical conditions and imaging modalities.

2. Related Work

2.1. Vision-Language Models

2.2. Prompt Learning



? 1. BiomedCoOp ?????????? LLM ??????? BiomedCLIP
??

??
[27]?????con-?????????????????????????? VLM
??????????????????? LLM ??????????????
GPT-4nostic ?? [17, 30]?????????????????????
VLM??? BiomedCLIP????????????????????VLM
??
??????????????ad-?????????????

2.3. Few-shot Adaptation of VLMs

?????????????els
??
ods??????????????CLIP-Adapter [16] ? Tip-Adapter [47]
??????????????????????????????????????CLIP-??????
??? MLP ???????Tip-Adapter
??
??????????????????????????? LP++
[19]??
????

sive hyperparameter optimization. Finally, CLAP [38] further constrains adaptation to stay close to original zero-shot prototypes by using adaptive penalties to prevent overfitting. These categories of methods refine visual embeddings typically at the final layers of VLMs, focusing on adjusting model features. In comparison, prompt learning approaches that target optimizing textual prompt inputs may be more advantageous in computational efficiency and adaptation to unseen classes, particularly in biomedical imaging domains.

3. Methodology

??? BiomedCoOp ???????? 1 ??????? BiomedCLIP
back-??
?????????????????????????????????(SCCM)
??
???????????????????-???????KDSP??????????
?????????????????????????????????SCCM ?
MSE ??? Kullback-Leibler (KL) ??KDSP
????????????????????

3.1. Contrastive Language-Image Pre-training

??? Xt ?RB×L ????????B????????????/?????H ? W ???????????????L ?

maxi-??????????????????-????????????????????? ?????????

? V = Ev(Xv) ? T = Et(Xt)??? V ?RB×D,T ?RB×C×D, C ???D

??

????? CLIP ?????????????????????????? C ?????????????????????? C ???????????????a photo of a [CLASS]?????aphoto of a????????????????? [CLASS] ?????????????? CoOp ?????????a ???????????????????????????????????????com-???????????????T(i)????? V ??????????

$$p(\boldsymbol{Y}=i|\boldsymbol{V},\boldsymbol{T}^{(i)})=\frac{\exp(\cos(\boldsymbol{T}^{(i)},\boldsymbol{V})/\tau)}{\sum_{j=1}^{C}\exp(\cos(\boldsymbol{T}^{(j)},\boldsymbol{V})/\tau)}, \quad (1)$$

where τ denotes the learnable temperature parameter and $\cos(\cdot,\cdot)$ denotes cosine similarity. The class of the image is determined by taking:

$$\arg\max_{\boldsymbol{i}} p(\boldsymbol{Y}=i|\boldsymbol{V},\boldsymbol{T}^{(i)}) \quad (2)$$

3.2. LLM Prompt Ensembling

???????????????????????[46]?????????[46]?????????[32]???[17, 30] ? GPT-4 [1] ???????????????????????????????? GPT-4 ?? LLM ???BiomedCoOp ???????????????????????????-???????????? C ????????????? Xg ?RN×C×L ?? N ???????????? Q ?????????? N ???????????????????

[CLASS] found in [MODALITY]." Unique for biomedical images, we specifically mention the imaging modality in Q as certain classes might overlap in different modalities. $\boldsymbol{X_g}$ from the LLM are then encoded to $\boldsymbol{T_g}=\boldsymbol{E_t}(\boldsymbol{X_g})\in\mathbb{R}^{N\times C\times D}$. For the SCCM component, all N text embeddings for each class are ensembled by taking the mean to get $\boldsymbol{P_g}\in\mathbb{R}^{C\times D}$:

$$- \sum_{i=1}^{N}\boldsymbol{T}_{\mathbf{g},\mathrm{i}}^{(c)} \quad (3)$$

3.3. Selective Prompting via Outlier Exclusion

?KDSP?????????????????con-?????????LLM ??-????????????????en-???????????????????????????????????for-?????????????????BiomedCLIP ?? B ??????????????????????????????????-??????? Ev ??????? V ????????? V ??????????Tg ??? S ??? tak- ????????????? logits ?????????

$$- \sum_{i=1}^{B}\max_{j}(\beta\cdot\boldsymbol{T}_{\boldsymbol{g}}^{(j)}\cdot\boldsymbol{V}^{\intercal}), \quad (4)$$

where β is a scaling factor applied to the logits.

To detect and handle anomalous prompts that deviate from the general distribution, we apply an outlier detection approach using the *Median Absolute Deviation* test statistic. Specifically, we calculate the median M_s of the prompt scores $\boldsymbol{S}$ and the median absolute deviation D:

For a given prompt, we compute the modified z-score:

$$z=\frac{\boldsymbol{S}-M_s}{D}. \quad (7)$$

We select only the N_s prompts that are associated with the modified z-scores with absolute values that are lower than a selection threshold ζ_s. Following a similar approach to Eq. 3, we obtain an average prompt encoding $\boldsymbol{P_s}$ based on the selections.

3.4. Overall Learning Objective

???????????????????ing
??
?????? Xp?????????? K ?????
-????????????????? logits ????????

$$\sum_{i=1}^{C} \boldsymbol{Y}^{(i)} \log p(\boldsymbol{Y} = i | \boldsymbol{V}, \boldsymbol{T_p}^{(i)}), \tag{8}$$

where $\boldsymbol{Y}^{(i)}$ is the ground truth label for class i, and $p(\boldsymbol{Y} = i|\boldsymbol{V}, \boldsymbol{T_p}^{(i)})$ is the predicted probability of class i given images $\boldsymbol{V}$ and encoded learnable text prompt $\boldsymbol{T_p}^{(i)}$.

Second, since the context to be learned is unified among all the classes, we minimize the difference between $\boldsymbol{P_g}$ and $\boldsymbol{T_p}$ in the SCCM component to ensure that the general biomedical knowledge is properly learned:

$$\sum_{i=1}^{C} ||\boldsymbol{T_p}^{(i)} - \boldsymbol{P_g}^{(i)}||_2^2 \tag{9}$$

Lastly, to align the distribution of the logits from image embeddings with learnable context prompts (student logits) and the logits from image embeddings with selective LLM-generated text embeddings (teacher logits), we minimize the KL divergence between these two distributions in the KDSP component:

$$- \qquad , \tag{10}$$

where $\boldsymbol{Pr}_{\text{teacher}}(i)$ is the probability distribution of the logits of $\boldsymbol{V}$ and $\boldsymbol{P_s}$, and $\boldsymbol{Pr}_{\text{student}}(i)$ is the probability distribution of the logits of $\boldsymbol{V}$ and $\boldsymbol{T_p}$.

The KL divergence term, $\mathcal{L}_{\texttt{KDSP}}$ restricts the model from drifting toward embeddings that are not representative of the actual medical scans. By minimizing this KL divergence, we guide the model to stay within a meaningful embedding space closely related to the content of the medical scans. This alignment helps ensure that the learned embeddings retain essential information about the

??????????????????????????

$$\mathcal{L} = \mathcal{L}_{\texttt{CE}} + \lambda_1 \mathcal{L}_{\texttt{SCCM}} + \lambda_2 \mathcal{L}_{\texttt{KDSP}} \tag{11}$$

where λ_1, and λ_2 are loss-balancing weights.

4. Experiments and Results

4.1. Experimental Setup

We assess the efficacy of our proposed BiomedCoOp framework across a comprehensive set of benchmark biomedical imaging datasets under multiple evaluation protocols designed to test accuracy and generalization within and across various few-shot image classification tasks.

Few-Shot Learning: To assess the model's performance under limited supervision, we conduct few-shot experiments with varying numbers of labeled examples per class (K = 1, 2, 4, 8, and 16 shots). This is critical for evaluating the model's ability to learn effectively from sparse data, a common scenario in biomedical applications, while retaining both task-specific and general domain knowledge.

Base-to-Novel Class Generalization: To assess model generalizability of our technique, each dataset is divided into base and novel classes. The model is trained on the base classes using a 16-shot setup and subsequently evaluated on both base and novel classes. This setup tests the model's ability to generalize to unseen classes within the same dataset, showcasing its potential to recognize novel disease presentations without additional fine-tuning.

Datasets: We conduct our experiments on 11 diverse medical imaging datasets covering 10 different organs and 9 imaging modalities: Computerized Tomography (CTKidney [20]), Dermatoscopy (DermaMNIST [10, 40]), Endoscopy (Kvasir [35]), Fundus Photography (RETINA [31, 36]), Histopathology (LC25000 [5], CHMNIST [22]), Magnetic Resonance Imaging (BTMRI [34]), Optical Coherence Tomography (OCTMNIST [23]), Ultrasound (BUSI [2]), and X-Ray (COVID-QU-Ex [39], KneeXray [7]). This selection includes complex datasets, such as brain MRI and ultrasound, ensuring a thorough assessment of the model's performance across a broad spectrum of biomedical imaging contexts. The detailed data split and tasks for the experiments are included in the *Supplementary Materials*.

Implementation details: We employed BiomedCLIP with a ViT-B/16 backbone, averaging results over three runs. The training was set to 100 epochs for few-shot and 50 epochs for base-to-novel benchmarking. We initialized the learnable context with the embedding vector corresponding to "`a photo of a`" and used 50 LLM prompts, a 0.0025 learning rate, a batch size of 4, and an SGD optimizer across datasets. Optimal values for λ_1, λ_2, and ζ_s were selected

? 1. ?????????????????? 11 ?????????? (%)????? 3 ????????? 3 ?????????±??????-??????????

Zero-shot Methods					
BiomedCLIP [48]			42.05		
BiomedCLIP [48] + Ensemble			52.27		
BiomedCLIP [48] + Selective Ensemble			53.72		
CLIP-based Adapter Methods					
CLIP-Adapter [16]	44.66 ± 2.97	43.91 ± 2.48	44.36 ± 1.94	45.42 ± 2.38	46.69 ± 1.71
Tip-Adapter [47]	49.19 ± 4.84	52.36 ± 6.57	57.33 ± 5.07	61.98 ± 5.76	67.15 ± 4.25
Tip-Adapter-F [47]	51.17 ± 8.33	52.74 ± 5.88	61.23 ± 6.22	65.91 ± 3.64	70.91 ± 2.65
Linear Probing Methods					
Standard LP [37]	47.25 ± 8.65	54.21 ± 7.80	61.00 ± 6.81	65.85 ± 4.89	69.40 ± 2.91
LP++ [19]	47.24 ± 7.68	53.18 ± 7.29	59.02 ± 6.93	63.69 ± 4.68	68.35 ± 3.59
Prompt Learning Methods					
CoOp [51]	50.16 ± 6.93	54.18 ± 4.31	59.75 ± 3.72	65.84 ± 3.66	69.62 ± 2.83
CoCoOp [50]	48.49 ± 4.39	51.28 ± 5.06	54.69 ± 4.79	61.08 ± 3.49	65.09 ± 2.87
KgCoOp [44]	50.85 ± 5.59	53.18 ± 4.33	57.82 ± 4.50	62.08 ± 2.59	62.84 ± 1.72
ProGrad [52]	51.88 ± 6.39	54.71 ± 4.46	60.42 ± 4.78	65.61 ± 3.02	67.13 ± 3.00
BiomedCoOp (Ours)	**57.03 ± 2.80**	**59.13 ± 3.64**	**63.95 ± 2.42**	**68.32 ± 2.65**	**72.42 ± 1.69**

??????????????????????????????????NVIDIA A100 GPU?40GB ??????

4.2. Few-shot Evaluation

??? BiomedCoOp ???????????????????????CoOp?CoCoOp?ProGrad?KgCoOp?????? CLIP ????????CLIP-Adapter?Tip-Adapter??????-F?????? ??????????????LP++?????????????????????????? MapLe?DCPL?????????????????????????????????pa ?????????????????????? LLM ?????? BiomedCLIP ??????????????? BiomedCLIP ???????????????????ble 1?BiomedCoOp ???????? -?????????????????????????? - ProGrad - ??? 5.2%?1 ?? 2 ????? 4.6%???????????? BiomedCoOp ?? ????????????? BiomedCoOp ??? K-shot ???????? BiomedCoOp effectiveness in prompt-based biomedical adaptation, ensuring reliable cross-dataset accuracy gains.

4.3. Base-to-Novel Generalization

????? BiomedCoOp ?????????????????????CoOp?CoCoOp?ProGrad ? KgCoOp???????????????????????????HM?????????? ?BUSI???????????????????????? 2 ????????BiomedCoOp ?????????? 5-?? CTKid ???????????????? 10%?????????? BiomedCoOp ??? ??? ?

4.4. Ablation Experiments

4.4.1. Effect of Different Components

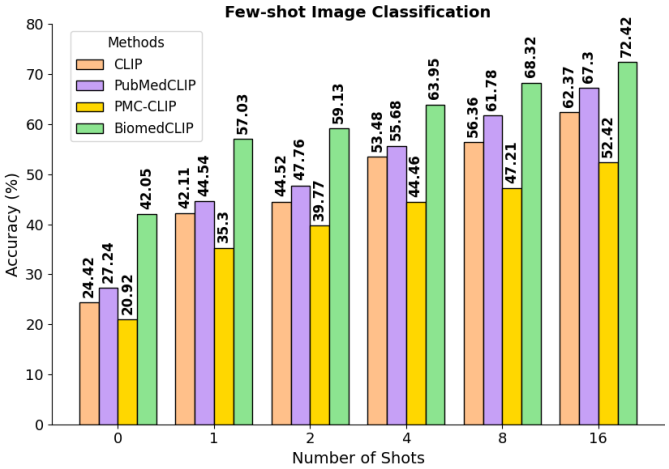
? 3 ????????????? BiomedCoOp ?????? BiomedCLIP???????? BiomedCoOp ??????????????????????????? SCCM ??????????????????????????? KDSP????????ac-?????????????????????????

Dataset		BiomedCLIP [48]	CoOp [51]	CoCoOp [50]	KgCoOp [43]	ProGrad [53]	BiomedCoOp (Ours)
	Base	47.84	73.85	72.26	68.36	71.67	**76.26**
	Novel	65.42	64.75	67.03	64.08	66.93	**73.92**
	HM	53.81	67.23	67.22	64.61	67.43	**75.07**
	Base	40.88	82.25	77.88	78.03	82.13	**82.42**
BTMRI	Novel	96.18	94.51	94.84	95.05	94.98	**96.84**
	HM	57.37	87.95	85.53	85.69	88.09	**89.05**
	Base	53.96	75.92	**77.28**	75.42	75.19	75.91
COVID-QU-Ex	Novel	89.43	90.07	87.61	89.61	90.34	**91.63**
	HM	67.31	82.39	82.12	81.90	82.07	**83.03**
	Base	38.55	82.24	81.96	81.67	83.86	**86.93**
CTKIDNEY	Novel	52.99	67.92	56.56	58.45	63.01	**78.94**
	HM	44.63	74.40	66.93	68.14	71.96	**82.74**
	Base	34.95	48.06	42.88	36.41	35.52	**54.86**
DermaMNIST	Novel	49.59	59.41	60.66	47.31	63.28	**74.1**
	HM	41.00	53.14	50.24	41.15	45.50	**63.04**
	Base	75.00	86.22	85.94	81.56	82.89	**86.50**
Kvasir	Novel	60.50	58.06	53.95	59.00	60.45	**61.83**
	HM	66.97	69.39	66.29	68.47	69.91	**72.11**
	Base	37.63	**89.41**	87.77	75.45	82.98	88.87
CHMNIST	Novel	40.69	35.11	42.51	38.70	**44.19**	42.73
	HM	39.10	50.42	57.28	51.16	57.67	**57.71**
	Base	59.73	90.12	88.33	88.13	90.29	**93.77**
LC25000	Novel	87.60	87.55	95.02	86.44	85.47	**97.00**
	HM	71.03	88.82	91.55	87.28	87.81	**95.36**
	Base	45.18	**70.98**	66.88	60.77	68.77	68.46
RETINA	Novel	55.28	56.90	65.56	54.91	58.43	**67.72**
	HM	49.72	63.16	66.21	57.69	63.18	**68.09**
	Base	35.89	38.28	34.08	37.94	40.88	**44.23**
KneeXray	Novel	71.90	47.69	63.14	61.19	59.12	**78.35**
	HM	47.88	42.47	44.27	46.84	48.34	**56.54**
	Base	56.60	75.00	79.6	68.20	74.20	**80.33**
OCTMNIST	Novel	50.00	50.23	**50.47**	50.13	50.02	50.07
	HM	53.10	60.17	**61.77**	57.79	59.76	61.69

? 2. ????????????(%)BiomedCoOp ? SOTA ????? HM= ??????????

Components			Base-to-Novel			Few-shot				
BiomedCLIP	SCCM	KDSP	Base	Novel	HM	1	2	4	8	16
✗	✗	✗	71.50	45.06	55.28	41.46	45.69	51.41	54.80	61.86
✗	✓	✗	71.91	43.04	53.85	43.72	44.77	52.76	60.08	64.04
✗	✗	✓	73.21	39.95	51.69	42.14	44.27	53.75	56.67	62.14
✗	✓	✓	72.95	39.04	50.86	42.11	44.52	53.48	56.36	62.37
✓	✗	✗	73.85	64.75	67.23	50.16	54.18	59.75	65.84	69.62
✓	✓	✗	75.21	65.79	70.19	51.62	54.99	61.43	65.93	69.99
✓	✗	✓	75.74	72.91	74.30	56.78	58.76	63.68	67.68	71.79
✓	✓	✓	**76.11**	**73.22**	**74.64**	**57.03**	**59.13**	**63.95**	**68.32**	**72.42**

? 3. ?? BiomedCoOp???????????? Base to Novel ?????? (%) ???????? BiomedCLIP ?? CLIP ?SCCM ? KDSP?



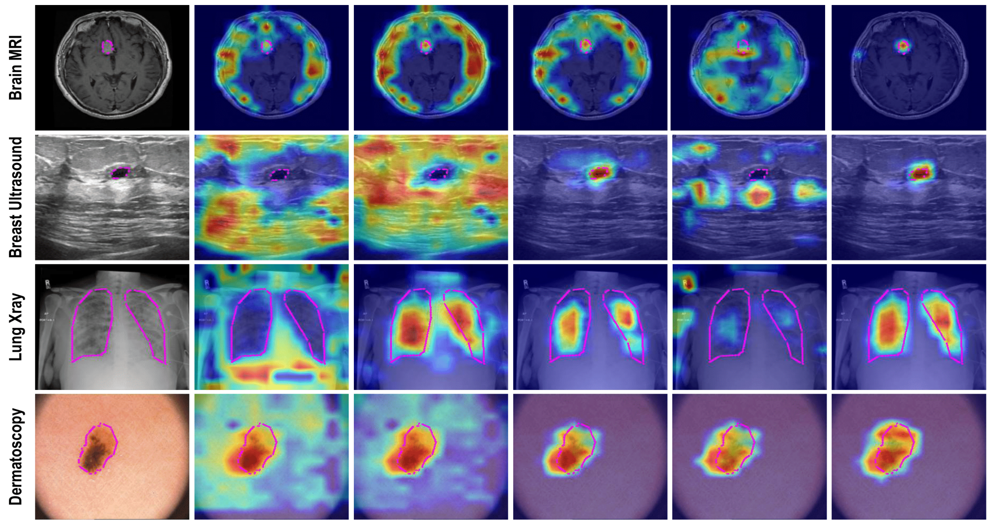
? 2. ???????? (%) ????? BiomedCOOP ? ????? CLIP ??????????????????

Prompts #	$K = 0$	$K = 1$	$K = 2$	$K = 4$	$K = 8$	$K = 16$
10	47.55	52.43	54.61	60.69	64.81	67.66
20	50.51	55.27	57.65	62.85	66.92	70.96
30	51.88	55.91	58.52	63.89	68.05	71.97
40	52.20	56.59	59.05	63.92	68.24	72.20
50	52.27	57.03	59.13	63.95	68.32	72.42

case-specific prompts. However, using KDSP with a CLIP-only setup hindered generalization, as the model lacked the domain knowledge needed to effectively exclude outliers (gray highlight). The combined use of SCCM and KDSP yielded optimal results (green highlight), balancing generalization and adaptability, especially in few-shot and base-to-novel tasks, affirming both components' essential roles in handling limited-data biomedical scenarios.

4.4.2. Effect of Number of LLM Prompts

Prompt diversity in relation to the number of LLM prompts may affect the quality of context mapping. To investigate this, Table 4 shows the effect of increasing LLM-generated prompts on BiomedCoOp's performance across few-shot settings (K = 0, 1, 2, 4, 8, and 16 shots). At lower shot levels (K = 0 and 1), a higher prompt count noticeably boosts accuracy, improving by 5-6% as prompts increase from 10 to 50. This indicates that prompt diversity is key for generalization with minimal labeled data. In intermediate shot



(a) Image (b) Manual Prompt (c) Prompt Ensemble (d) CoOp (e) CoCoOp (f) BiomedCoOp

low-data scenarios.

4.4.3. Effect of Different CLIP-based Models

????????? CLIP ?????????? BiomedCoOp ??????????-??????? CLIP ????????? CLIP (ViT-B/16)?PubMedCLIP (ViT-B/32)?PMC-CLIP (RN50) ? BiomedCLIP (ViT-B/16)?????? 2 ????????????? CLIP ??? BiomedCoOp ?????????BiomedCLIP ???????ac-?????????????????????????????????? ???????????????????BiomedCLIP ? 16 ?????? 72.42% ?????????????? CLIP ? PMC-CLIP ????????? PubMedCLIP ????????????? BiomedCLIP?????????????BiomedCLIP ?? VLM ????????????? VLM ???????????????????????

4.4.4. Visual Interpretability

??????gScoreCAM [8] ????????????????????????????????3?????????dis-Tinct ???????????????????????????[CLASS] ???????????????????????????????????????c??av-erages??LLM??????????-???????? ??????????? CoOp?CoCoOp ? BiomedCoOp?d?e?f ?????????????????

??[CLASS] ???? BUSI??????????COVID-QU-Ex??? X ??????? [9]?? ISIC [11,12,41]?????????BiomedCoOp?f ?????????????????????? giions????????????????? MRI ?? ???[28, 29]?????-??????????????????????????????????

5. Conclusion

????? BiomedCoOp ?? ?????????????????????????CoOp ?? ??????????????????????ment ?????????????????????????????????? BiomedCLIP ?? VLM ???????????????? ???

[1] Josh Achiam, Steven Adler, Sandhini Agarwal, Lama Ahmad, Ilge Akkaya, Florencia Leoni Aleman, Diogo Almeida, Janko Altenschmidt, Sam Altman, Shyamal Anadkat, et al. Gpt-4 technical report. *arXiv preprint arXiv:2303.08774*, 2023. 4

[2] Walid Al-Dhabyani, Mohammed Gomaa, Hussien Khaled, and Aly Fahmy. Dataset of breast ultrasound images. *Data in brief*, 28:104863, 2020. 5, 16

[3] Yequan Bie, Luyang Luo, Zhixuan Chen, and Hao Chen. Xcoop: Explainable prompt learning for computer-aided diagnosis via concept-guided context optimization. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 773–783. Springer, 2024. 2, 13

[4] Benedikt Boecking, Naoto Usuyama, Shruthi Bannur, Daniel C Castro, Anton Schwaighofer, Stephanie Hyland, Maria Wetscherek, Tristan Naumann, Aditya Nori, Javier Alvarez-Valle, et al. Making the most of text semantics to improve biomedical vision–language processing. In *European conference on computer vision*, pages 1–21. Springer, 2022. 2

[5] Andrew A. Borkowski, Marilyn M. Bui, L. Brannon Thomas, Catherine P. Wilson, Lauren A. DeLand, and Stephen M. Mastorides. Lung and colon cancer histopathological image dataset (lc25000), 2019. 5, 16

[6] Qinglong Cao, Zhengqin Xu, Yuntian Chen, Chao Ma, and Xiaokang Yang. Domain-controlled prompt learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 936–944, 2024. 2, 13

[7] Pingjun Chen. Knee osteoarthritis severity grading dataset, 2018. 5, 16

[8] Peijie Chen, Qi Li, Saad Biaz, Trung Bui, and Anh Nguyen. gscorecam: What objects is clip looking at? In *Proceedings of the Asian Conference on Computer Vision*, pages 1959–1975, 2022. 8

[9] Jun Cheng. brain tumor dataset, 2017. 8

[10] Noel Codella, Veronica Rotemberg, Philipp Tschandl, M Emre Celebi, Stephen Dusza, David Gutman, Brian Helba, Aadi Kalloo, Konstantinos Liopyris, Michael Marchetti, et al. Skin lesion analysis toward melanoma detection 2018: A challenge hosted by the international skin imaging collaboration (isic). *arXiv preprint arXiv:1902.03368*, 2019. 5, 16

[11] Noel CF Codella, David Gutman, M Emre Celebi, Brian Helba, Michael A Marchetti, Stephen W Dusza, Aadi Kalloo, Konstantinos Liopyris, Nabin Mishra, Harald Kittler, et al. Skin lesion analysis toward melanoma detection: A challenge at the 2017 international symposium on biomedical imaging (isbi), hosted by the international skin imaging collaboration (isic). In *2018 IEEE 15th international symposium on biomedical imaging (ISBI 2018)*, pages 168–172. IEEE, 2018. 8

[12] Marc Combalia, Noel CF Codella, Veronica Rotemberg, Brian Helba, Veronica Vilaplana, Ofer Reiter, Cristina Carrera, Alicia Barreiro, Allan C Halpern, Susana Puig, et al. Bcn20000: Dermoscopic lesions in the wild. *arXiv preprint arXiv:1908.02288*, 2019. 8

[13] Yuexi Du, Brian Chang, and Nicha C Dvornek. Cleft: Language-image contrastive learning with efficient large language model and prompt fine-tuning. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 465–475. Springer, 2024. 2

[14] Sedigheh Eslami, Gerard de Melo, and Christoph Meinel. Does clip benefit visual question answering in the medical domain as much as it does in the general domain?, 2021. 2

[15] Xiao Fang, Yi Lin, Dong Zhang, Kwang-Ting Cheng, and Hao Chen. Aligning medical images with general knowledge from large language models. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 57–67. Springer, 2024. 2

[16] Peng Gao, Shijie Geng, Renrui Zhang, Teli Ma, Rongyao Fang, Yongfeng Zhang, Hongsheng Li, and Yu Qiao. Clip-adapter: Better vision-language models with feature adapters. *International Journal of Computer Vision*, 132(2): 581–595, 2024. 1, 3, 6

[17] Roman Johannes Gertz, Thomas Dratsch, Alexander Christian Bunck, Simon Lennartz, Andra-Iza Iuga, Martin Gunnar Hellmich, Thorsten Persigehl, Lenhard Pennig, Carsten Herbert Gietzen, Philipp Fervers, et al. Potential of gpt-4 for detecting errors in radiology reports: Implications for reporting accuracy. *Radiology*, 311(1):e232714, 2024. 3, 4

[18] Neil Houlsby, Andrei Giurgiu, Stanislaw Jastrzebski, Bruna Morrone, Quentin De Laroussilhe, Andrea Gesmundo, Mona Attariyan, and Sylvain Gelly. Parameter-efficient transfer learning for nlp. In *International conference on machine learning*, pages 2790–2799. PMLR, 2019. 1

[19] Yunshi Huang, Fereshteh Shakeri, Jose Dolz, Malik Boudiaf, Houda Bahig, and Ismail Ben Ayed. Lp++: A surprisingly strong linear probe for few-shot clip. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 23773–23782, 2024. 1, 3, 6

[20] Md Nazmul Islam, Mehedi Hasan, Md Kabir Hossain, Md Golam Rabiul Alam, Md Zia Uddin, and Ahmet Soylu. Vision transformer and explainable transfer learning models for auto detection of kidney cyst, stone and tumor from ct-radiography. *Scientific Reports*, 12(1):1–14, 2022. 5, 16

[21] Chao Jia, Yinfei Yang, Ye Xia, Yi-Ting Chen, Zarana Parekh, Hieu Pham, Quoc Le, Yun-Hsuan Sung, Zhen Li, and Tom Duerig. Scaling up visual and vision-language representation learning with noisy text supervision. In *International conference on machine learning*, pages 4904–4916. PMLR, 2021. 2

[22] Jakob Nikolas Kather, Cleo-Aron Weis, Francesco Bianconi, Susanne M Melchers, Lothar R Schad, Timo Gaiser, Alexander Marx, and Frank Gerrit Zöllner. Multi-class texture analysis in colorectal cancer histology. *Scientific reports*, 6(1): 1–11, 2016. 5, 16

[23] Daniel S. Kermany, Michael Goldbaum, et al. Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell*, 172(5):1122 – 1131.e9, 2018. 5, 16

[24] Muhammad Uzair Khattak, Hanoona Rasheed, Muhammad Maaz, Salman Khan, and Fahad Shahbaz Khan. Maple:

Multi-modal prompt learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 19113–19122, 2023. 2

[25] Muhammad Uzair khattak, Hanoona Rasheed, Muhammad Maaz, Salman Khan, and Fahad Shahbaz Khan. Maple: Multi-modal prompt learning. In *The IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023. 1

[26] Muhammad Uzair Khattak, Syed Talal Wasim, Muzammal Naseer, Salman Khan, Ming-Hsuan Yang, and Fahad Shahbaz Khan. Self-regulating prompts: Foundational model adaptation without forgetting. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 15190–15200, 2023. 2

[27] Muhammad Uzair Khattak, Muhammad Ferjad Naeem, Muzammal Naseer, Luc Van Gool, and Federico Tombari. Learning to prompt with text only supervision for vision-language models. *arXiv preprint arXiv:2401.02418*, 2024. 2, 3

[28] Taha Koleilat, Hojat Asgariandehkordi, Hassan Rivaz, and Yiming Xiao. Medclip-sam: Bridging text and image towards universal medical image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 643–653. Springer, 2024. 8

[29] Taha Koleilat, Hojat Asgariandehkordi, Hassan Rivaz, and Yiming Xiao. Medclip-samv2: Towards universal text-driven medical image segmentation. *arXiv preprint arXiv:2409.19483*, 2024. 8

[30] Satheesh Krishna, Nishaant Bhambra, Robert Bleakney, and Rajesh Bhayana. Evaluation of reliability, repeatability, robustness, and confidence of gpt-3.5 and gpt-4 on a radiology board–style examination. *Radiology*, 311(2):e232715, 2024. 3, 4

[31] Thomas Köhler, Attila Budai, Martin Kraus, Jan Odstrcilik, Georg Michelson, and Joachim Hornegger. Automatic no-reference quality assessment for retinal fundus images using vessel segmentation, 2013. 5, 16

[32] Jiaxiang Liu, Tianxiang Hu, Yan Zhang, Xiaotang Gai, Yang Feng, and Zuozhu Liu. A chatgpt aided explainable framework for zero-shot medical image diagnosis. *arXiv preprint arXiv:2307.01981*, 2023. 4

[33] Jun Ma, Yuting He, Feifei Li, Lin Han, Chenyu You, and Bo Wang. Segment anything in medical images. *Nature Communications*, 15(1):654, 2024. 2

[34] Msoud Nickparvar. Brain tumor mri dataset, 2021. 5, 16

[35] Konstantin Pogorelov, Kristin Ranheim Randel, Carsten Griwodz, Sigrun Losada Eskeland, Thomas de Lange, Dag Johansen, Concetto Spampinato, Duc-Tien Dang-Nguyen, Mathias Lux, Peter Thelin Schmidt, Michael Riegler, and Pål Halvorsen. Kvasir: A multi-class image dataset for computer aided gastrointestinal disease detection. In *Proceedings of the 8th ACM on Multimedia Systems Conference*, pages 164–169, New York, NY, USA, 2017. ACM. 5, 16

[36] Prasanna Porwal, Samiksha Pachade, Ravi Kamble, Manesh Kokare, Girish Deshmukh, Vivek Sahasrabuddhe, and Fabrice Meriaudeau. Indian diabetic retinopathy image dataset (idrid), 2018. 5, 16

[37] Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, et al. Learning transferable visual models from natural language supervision. In *International conference on machine learning*, pages 8748–8763. PMLR, 2021. 1, 2, 6

[38] Julio Silva-Rodriguez, Sina Hajimiri, Ismail Ben Ayed, and Jose Dolz. A closer look at the few-shot adaptation of large vision-language models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 23681–23690, 2024. 3

[39] Anas M. Tahir, Muhammad E.H. Chowdhury, Amith Khandakar, Tawsifur Rahman, Yazan Qiblawey, Uzair Khurshid, Serkan Kiranyaz, Nabil Ibtehaz, M. Sohel Rahman, Somaya Al-Maadeed, Sakib Mahmud, Maymouna Ezeddin, Khaled Hameed, and Tahir Hamid. Covid-19 infection localization and severity grading from chest x-ray images. *Computers in Biology and Medicine*, 139:105002, 2021. 5, 16

[40] Philipp Tschandl, Cliff Rosendahl, and Harald Kittler. The ham10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Scientific data*, page 180161, 2018. 5, 16

[41] Philipp Tschandl, Cliff Rosendahl, and Harald Kittler. The ham10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Scientific data*, 5(1):1–9, 2018. 8

[42] Yan Xu, Rixiang Quan, Weiting Xu, Yi Huang, Xiaolong Chen, and Fengyuan Liu. Advances in medical image segmentation: A comprehensive review of traditional, deep learning and hybrid approaches. *Bioengineering*, 11(10): 1034, 2024. 2

[43] Hantao Yao, Rui Zhang, and Changsheng Xu. Visual-language prompt tuning with knowledge-guided context optimization, 2023. 2, 7

[44] Hantao Yao, Rui Zhang, and Changsheng Xu. Visual-language prompt tuning with knowledge-guided context optimization. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 6757–6767, 2023. 6

[45] Jamil Zaghir, Marco Naguib, Mina Bjelogrlic, Aurélie Névéol, Xavier Tannier, and Christian Lovis. Prompt engineering paradigms for medical applications: scoping review and recommendations for better practices. *arXiv preprint arXiv:2405.01249*, 2024. 7

[46] Chenrui Zhang, Lin Liu, Chuyuan Wang, Xiao Sun, Hongyu Wang, Jinpeng Wang, and Mingchen Cai. Prefer: Prompt ensemble learning via feedback-reflect-refine. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 19525–19532, 2024. 4

[47] Renrui Zhang, Rongyao Fang, Wei Zhang, Peng Gao, Kunchang Li, Jifeng Dai, Yu Qiao, and Hongsheng Li. Tip-adapter: Training-free clip-adapter for better vision-language modeling. *arXiv preprint arXiv:2111.03930*, 2021. 1, 3, 6

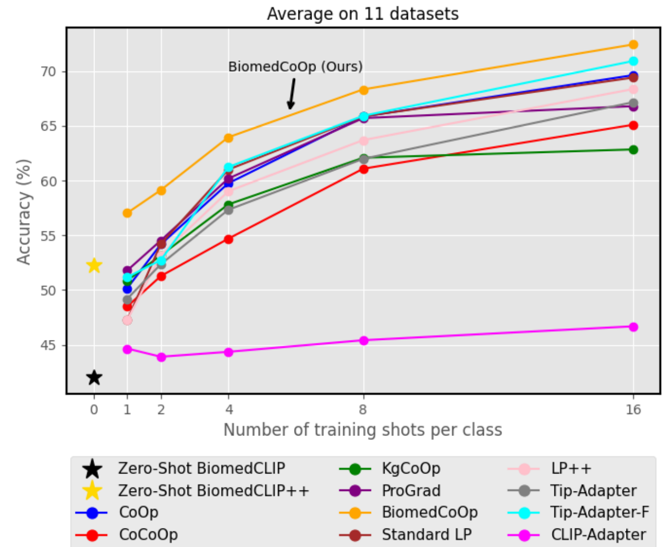
[48] Sheng Zhang, Yanbo Xu, Naoto Usuyama, Hanwen Xu, Jaspreet Bagga, Robert Tinn, Sam Preston, Rajesh Rao, Mu Wei, Naveen Valluri, Cliff Wong, Andrea Tupini, Yu

Wang?Matt Mazzola?Swadheen Shukla?Lars Liden?Jian-??Matthew P. Lungren?Tristan Naumann??????.????????????????????? 1500 ?????????2024. 1, 2, 6, 7[49]
222 arXiv:2312.07353, 2023. 2[50] Kaiyang Zhou?Jingkang Yang?Chen Change Loy?Zi-??.????????????? CVPR?? 16816?16825 ??2022
??1,2,6,7[51]???????????????.??????????? IJCV,130(9):2337?2348, 2022. 1, 2, 6, 7[52]?????????????????-??.???????????????IEEE/CVF ????????????? 15659?15669 ??2023 ??
6[53]?????????????????.???????????????2024?2??7

Wang?Matt Mazzola?Swadheen Shukla?Lars Liden?Jian-??Matthew P. Lungren?Tristan Naumann??????.????????????????????? 1500 ?????????2024. 1, 2, 6, 7[49]
222 arXiv:2312.07353, 2023. 2[50] Kaiyang Zhou?Jingkang Yang?Chen Change Loy?Zi-??.????????????? CVPR?? 16816?16825 ??2022
??1,2,6,7[51]???????????????.??????????? IJCV,130(9):2337?2348, 2022. 1, 2, 6, 7[52]?????????????????-??.???????????????IEEE/CVF ????????????? 15659?15669 ??2023 ??
6[53]?????????????????.???????????????2024?2??7

6. Detailed Dataset Overview

?S7???????11 ????????????BiomedCoOp???9??????????????? CT?MRI?X ?????????10???/??/??????????????????? CT ?????????????????????????????X ??? ??? ???????????????????? ples ????????????????????????????????????-BiomedCoOp ???????????



?S1????????????????%??????????????????????????

7. Additional Few-shot Results

?S1???????BiomedCoOp ??????????????????K = 1,2,4,8,16?????BiomedCoOp ????????????????????????????????-? S8 ?????????????????????-CoOp ? SOTA ??????????? SOTA ?? -????????????

8. Learnable Context Interpretability

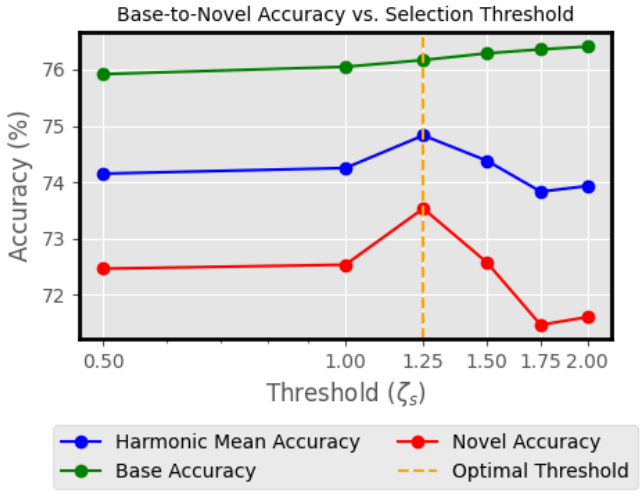
??[??]???????S9?? Kvasir ??endoscopy??BTMRI ??mri?????????????????????-?????????????

9. Effect of Context Length

As shown in Table S1, increasing the context length tends to reduce performance on both base and novel classes. A shorter context length, such as 4, achieves a better balance between base and novel accuracy, resulting in a higher harmonic mean (HM) score. As the context length increases to 16, 32, and 64, the accuracy on novel classes declines more rapidly than on base classes, leading to a sharp reduction in the harmonic mean. This pattern suggests that longer context lengths could diminish the model's ability to generalize effectively across both base and novel classes.

Context Length	Base Acc.	Novel Acc.	HM
4	76.11	73.22	74.64
16	74.93	67.98	71.29
32	72.34	62.73	67.19
64	71.50	58.99	64.65

? S1??????????????acc?????????????????????%???



?S2?????((?s)??Base-to-Novel??????

10. Effect of Prompt Selection Threshold

????????????KDSP??????????????????????????
?? (?s) ????? z ????????????? LLM
????????S???.????????(?s) ????? z
??????????????LLM??????????????????????????s=1.25??????
??
??
??????????????????????????s=1.25????????????
????????????????????

11. Selective Prompting for SCCM

??????Con-????????SCCM????sistenc
?????????????
y??? BiomedCoOp ???????????
SCCM?????????????????S5????????????????????????76.2
SCCM?????????????????S5????????????????????????
?73.92%????????????SCCM ????????????????????????????
??????????????76.26%76.39%????????????73
.92% ?72.59%??????????74.44%
?73.92%????????SCCM
????????????????????????????

alization to novel classes. Thus, keeping all prompt samples in the mapping process (i.e., SCCM) helps maintain broader generalization, balancing performance across both base and novel classes.

12. Additional Comparisons with Other Recent Methods

??????????? SOTA ????? CLIP ?????XCoOp [3] ? DCPL
[6]??S2?S3??
???????? 16 ????? S3 ??????????????????DCPL????
[6]?L-S2? BiomedCLIP ??LSDM ??????????XCoOp????
BiomedCLIP
???????????????????? 16 ????? S3
????????????????????????????LLM??????????????????????
?????????????????????DCPL???? DCPL*????
????? S2??
BiomedCLIP ??LSDM ??????????? XCoOp????
BiomedCLIP
????????????????????????????LLM???????
????????????????????? S2??

?S2????????(%) ???SOTA??????????????DCPL*???
BiomedCLIP ??LSDM ??????????XCoOp????BiomedCLIP
??????????????????

BiomedCLIP			42.05		
BiomedCLIP + Ensemble			52.27 (+10.22 from BiomedCLIP)		
BiomedCLIP + Selection			53.72 (+11.67 from BiomedCLIP)		
DCPL*	$45.65_{8.86}$	$51.65_{8.79}$	$56.62_{7.51}$	$62.85_{8.40}$	$68.79_{4.80}$
XCoOp	$52.50_{5.91}$	$55.39_{5.74}$	$60.87_{4.18}$	$66.37_{3.44}$	$71.04_{1.95}$
BiomedCoOp (Ours)	$57.03_{2.80}$	$59.13_{3.64}$	$63.95_{2.42}$	$68.32_{2.65}$	$72.42_{1.69}$

? S3??? SOTA ??????????????DCPL*??BiomedCLIP
??????? LSDM??XCoOp????BiomedCLIP????

XCoOp	74.62%	63.19%	68.43%
DCPL*	68.70%	40.35%	50.84%
BiomedCoOp (Ours)	**76.26%**	**73.92%**	**75.07%**

13. Effect of LLM used

?S4?????50????????4-shot
??Gemma-2-2b???
??
??????Gemma-2-2b???????????????????????

Method	BTMRI	COVID-QU-Ex	CTKidney
LLaMA-3-8b	$76.61_{3.53}$	$73.20_{1.84}$	$\mathbf{67.81_{0.39}}$
Gemma-2-2b	$76.69_{3.46}$	$72.46_{2.46}$	$66.03_{0.96}$
GPT-4	$\mathbf{77.23_{3.90}}$	$\mathbf{73.28_{2.30}}$	$66.50_{1.92}$

Component	Base Acc.	Novel Acc.	HM
SCCM without SPOE	76.26	73.92	75.07
SCCM with SPOE	76.39	72.59	74.44

14. Additional Hyperparameters

? S6 ????????????1?????? ?s) ?? BiomedCoOp ??1??2????????????????????? ? ?????????????????? zs ?????????[1.25, 1.5] ??????????????????????????

? S6 ????????????1??2?? ?s) ?? BiomedCoOp ???1 ??2?????????????????? ????????????????? zs ????????[1.25, 1.5] ????????????????????????

BTMRI	Few-shot	0.5	0.25	1.5
	Base-to-Novel	0.5	0.5	1.25
BUSI	Few-shot	0.75	0.75	1.5
	Base-to-Novel	-	-	-
COVID-QU-Ex	Few-shot	0.5	2.0	1.5
	Base-to-Novel	20.0	1.0	1.25
CTKIDNEY	Few-shot	1.0	0.5	1.5
	Base-to-Novel	10.0	0.25	1.25
DermaMNIST	Few-shot	5.0	20.0	1.5
	Base-to-Novel	2.0	0.5	1.5
Kvasir	Few-shot	0.75	0.75	1.5
	Base-to-Novel	1.0	1.0	1.25
CHMNIST	Few-shot	0.25	0.25	1.5
	Base-to-Novel	10.0	1.0	1.5
LC25000	Few-shot	0.5	0.5	1.5
	Base-to-Novel	0.25	0.75	1.25
RETINA	Few-shot	0.25	0.25	1.5
	Base-to-Novel	5.0	1.0	2.0
KneeXray	Few-shot	5.0	20.0	1.75
	Base-to-Novel	0.25	3.0	1.25
OCTMNIST	Few-shot	1.0	0.75	1.5
	Base-to-Novel	0.75	0.5	1.5

? S6?????? ?1??2???s????????????????????????

15. LLM Prompts Used

??????????? GPT-4???????????????????????????

“The image of a normal brain on MRI shows a clear differentiation between different brain regions with no disruptions.”

??????????MRI ?????????????

“Meningioma tumor on MRI displaying a dural tail sign and homogeneous enhancement.”

“Pituitary tumors often cause sellar expansion and may invade adjacent structures.”

“A routine ultrasound showing a hypoechoic, well-defined nodule, indicating a benign breast tumor.”

“An ultrasound revealing microcalcifications within the mass, indicating a malignant breast tumor.”

“A grayscale ultrasound highlighting well-defined ducts and lobules, characteristic of a normal breast ultrasound scan.”

“An X-ray scan showing bilateral airspace consolidation, typical of covid lungs.”

“A chest X-ray image with reticular and nodular opacities, indicative of lung opacity lungs.”

“An X-ray revealing no signs of consolidation or effusion, suggesting normal lungs.”

"An X-ray image revealing multifocal ground-glass and consolidative opacities, indicative of viral pneumonia lungs."

"A CT image showing a lesion with uniform density and no internal irregularities, indicative of a cyst kidney."

"A CT scan showing a calcified structure with acoustic shadowing, consistent with a kidney stone."

"A CT scan showing a lesion with poorly defined margins, consistent with a kidney tumor."

"A CT image revealing no signs of renal atrophy or cortical thinning, suggesting a normal kidney."

"Actinic keratosis lesions may become thicker and more pronounced over time without treatment."

"BCC lesions may bleed with minor trauma, such as shaving, due to their friable nature."

"Cryotherapy, using liquid nitrogen, is a common treatment for seborrheic keratosis, causing the lesions to blister and fall off."

"Dermatofibromas can be multiple in patients with systemic lupus erythematosus or other autoimmune conditions."

"A clinical image with a lesion that has changed in size or texture, indicative of melanoma."

"Melanocytic nevi can become darker and larger during pregnancy due to hormonal changes and increased melanin production."

"The diagnosis of vascular lesions often requires a combination of clinical examination and sometimes imaging studies."

"Dyed lifted polyps can exhibit various morphological features, including lobulated, sessile, or pedunculated appearances."

"Endoscopic images of dyed resection margins often show a bright, distinct color outlining the area of resection, contrasting with the surrounding mucosa."

"In severe cases, esophagitis may lead to strictures or narrowing of the esophageal lumen, visible during endoscopy."

"Endoscopic images of the normal cecum show a well-defined junction with the ascending colon, without any transitional abnormalities."

"Endoscopic examination of the normal pylorus shows a lack of any masses, polyps, or other abnormal growths."

"The Z line in a normal endoscopy appears intact and well-defined, with no evidence of structural compromise."

"Polyps can be classified based on their appearance and histological features, including adenomatous polyps, hyperplastic polyps, or inflammatory polyps."

"Ulcerative colitis can be associated with extra-intestinal manifestations, including dermatological, joint, ocular, or hepatobiliary complications."

Modality	Organ(s)	Name	Classes	# train/val/test
Computerized Tomography	Kidney	CTKidney [20]	Kidney Cyst, Kidney Stone, Kidney Tumor, Normal Kidney	6221/2487/3738
Dermatoscopy	Skin	DermaMNIST [10, 40]	Actinic Keratosis, Basal Cell Carcinoma, Benign Keratosis, Dermatofibroma, Melanocytic nevus, Melanoma, Vascular Lesion	7007/1003/2005
			Dyed Lifted Polyps, Normal Cecum, Esophagitis, Dyed Resection Margins, Normal Pylorus, Normal Z Line, Polyps, Ulcerative Colitis	2000/800/1200
Fundus Photography	Retina	RETINA [31, 36]	Cataract, Diabetic Retinopathy, Glaucoma, Normal Retina	2108/841/1268
Histopathology		???LC25000 [5]	Colon Adenocarcinoma, Colon Benign Tissue, Lung Adenocarcinoma, Lung Benign Tissue, Lung Squamous Cell Carcinoma	12500/5000//7500
			Adipose Tissue, Complex Stroma, Debris, Empty Background, Immune Cells, Normal Mucosal Glands, Simple Stroma, Tumor Epithelium	2496/1000/1504
Magnetic Resonance Imaging	Brain	BTMRI [34]	Glioma Tumor, Meningioma Tumor, Normal Brain, Pituitary Tumor	2854/1141/1717
Optical Coherence Tomography	Retina	OCTMNIST [23]	Choroidal Neovascularization, Drusen, Diabetic Macular Edema, Normal	97477/10832/1000
Ultrasound	Breast	BUSI [2]	Benign Tumors, Malignant Tumors, Normal Scans	389/155/236
X-Ray	Chest	COVID-QU-Ex [39]	COVID-19, Lung Opacity, Normal Lungs, Viral Pneumonia	10582/4232/6351
	Knee	KneeXray [7]	No, Doubtful, Minimal, Moderate, and Severe Osteoarthritis	5778/826/1656

BTMRI	BiomedCLIP			56.79		
	BiomedCLIP + Ensemble			61.04		
	CLIP-Adapter	$56.80_{\pm 0.48}$	$57.13_{\pm 0.88}$	$56.80_{\pm 0.48}$	$57.15_{\pm 0.91}$	$60.16_{\pm 0.32}$
	Tip-Adapter	$66.66_{\pm 4.37}$	$67.77_{\pm 2.74}$	$76.37_{\pm 1.69}$	$73.75_{\pm 3.15}$	$78.97_{\pm 1.25}$
	Tip-Adapter-F	$59.60_{\pm 2.28}$	$61.94_{\pm 6.74}$	$77.90_{\pm 1.71}$	$79.18_{\pm 1.80}$	$82.27_{\pm 2.33}$
	Standard LP	$62.24_{\pm 5.03}$	$72.45_{\pm 5.27}$	$75.98_{\pm 1.94}$	$77.63_{\pm 3.45}$	$81.24_{\pm 2.56}$
	LP++	$64.72_{\pm 6.16}$	$71.69_{\pm 5.88}$	$75.48_{\pm 1.41}$	$77.11_{\pm 1.28}$	$81.61_{\pm 1.31}$
	CoOp	$63.82_{\pm 3.94}$	$68.82_{\pm 5.15}$	$74.68_{\pm 2.99}$	$79.27_{\pm 1.9}$	$82.37_{\pm 1.89}$
	CoCoOp	$59.47_{\pm 0.78}$	$64.14_{\pm 0.64}$	$67.83_{\pm 4.8}$	$71.69_{\pm 4.4}$	$78.45_{\pm 1.83}$
	KgCoOp	$63.33_{\pm 3.66}$	$70.16_{\pm 5.47}$	$75.4_{\pm 2.45}$	$79.79_{\pm 0.99}$	$81.07_{\pm 0.33}$
	ProGrad	$66.92_{\pm 2.10}$	$71.46_{\pm 3.46}$	$76.24_{\pm 5.07}$	$78.82_{\pm 1.77}$	$82.84_{\pm 1.02}$
	BiomedCoOp (Ours)	$65.08_{\pm 1.81}$	$70.57_{\pm 4.31}$	$77.23_{\pm 3.9}$	$78.55_{\pm 2.19}$	$83.3_{\pm 1.34}$
BUSI	BiomedCLIP			59.75		
	BiomedCLIP + Ensemble			59.75		
	CLIP-Adapter	$61.44_{\pm 0.78}$	$61.01_{\pm 1.03}$	$61.72_{\pm 0.81}$	$61.86_{\pm 1.41}$	$63.55_{\pm 2.17}$
	Tip-Adapter	$62.71_{\pm 2.56}$	$61.44_{\pm 2.44}$	$59.03_{\pm 1.13}$	$55.93_{\pm 11.37}$	$68.78_{\pm 5.54}$
	Tip-Adapter-F	$61.86_{\pm 2.17}$	$56.35_{\pm 7.25}$	$64.54_{\pm 7.01}$	$68.50_{\pm 2.26}$	$71.89_{\pm 1.25}$
	Standard LP	$51.41_{\pm 10.78}$	$47.88_{\pm 6.44}$	$53.38_{\pm 7.12}$	$65.53_{\pm 6.34}$	$68.78_{\pm 1.80}$
	LP++	$51.12_{\pm 4.95}$	$55.50_{\pm 2.38}$	$60.31_{\pm 3.42}$	$66.10_{\pm 2.34}$	$70.05_{\pm 1.58}$
	CoOp	$48.73_{\pm 3.3}$	$53.53_{\pm 2.8}$	$60.17_{\pm 3.65}$	$64.69_{\pm 6.4}$	$69.49_{\pm 3.3}$
	CoCoOp	$52.26_{\pm 3.73}$	$49.15_{\pm 2.77}$	$59.75_{\pm 1.83}$	$65.82_{\pm 3.83}$	$70.2_{\pm 1.22}$
	KgCoOp	$53.39_{\pm 7.25}$	$55.51_{\pm 3.30}$	$62.01_{\pm 4.38}$	$67.37_{\pm 2.42}$	$70.62_{\pm 2.11}$
	ProGrad	$46.33_{\pm 4.23}$	$49.15_{\pm 7.32}$	$62.29_{\pm 7.49}$	$64.83_{\pm 4.20}$	$71.47_{\pm 2.69}$
	BiomedCoOp (Ours)	$50.71_{\pm 1.74}$	$50.71_{\pm 7.34}$	$59.32_{\pm 1.04}$	$63.27_{\pm 4.61}$	$70.34_{\pm 2.27}$
COVID-QU-Ex	BiomedCLIP			43.8		
	BiomedCLIP + Ensemble			66.86		
	CLIP-Adapter	$50.42_{\pm 1.55}$	$43.04_{\pm 1.16}$	$46.28_{\pm 3.30}$	$48.68_{\pm 1.13}$	$49.55_{\pm 1.35}$
	Tip-Adapter	$62.13_{\pm 7.82}$	$58.72_{\pm 5.19}$	$63.84_{\pm 10.41}$	$66.77_{\pm 5.64}$	$73.05_{\pm 1.04}$
	Tip-Adapter-F	$54.89_{\pm 17.51}$	$54.01_{\pm 7.87}$	$69.97_{\pm 4.13}$	$69.89_{\pm 4.08}$	$76.07_{\pm 3.22}$
	Standard LP	$49.91_{\pm 10.98}$	$48.06_{\pm 16.94}$	$60.55_{\pm 13.60}$	$68.29_{\pm 6.12}$	$71.98_{\pm 1.88}$
	LP++	$46.41_{\pm 10.75}$	$56.42_{\pm 15.04}$	$62.32_{\pm 9.54}$	$66.19_{\pm 8.40}$	$72.79_{\pm 1.17}$
	CoOp	$58.82_{\pm 14.51}$	$58.37_{\pm 8.14}$	$67.03_{\pm 6.58}$	$74.66_{\pm 0.29}$	$76.37_{\pm 1.39}$
	CoCoOp	$69.36_{\pm 2.79}$	$68.8_{\pm 2.65}$	$63.7_{\pm 10.27}$	$69.36_{\pm 3.28}$	$74.52_{\pm 0.72}$
	KgCoOp	$61.68_{\pm 9.84}$	$54.68_{\pm 12.19}$	$65.91_{\pm 8.61}$	$74.86_{\pm 0.28}$	$75.65_{\pm 0.88}$
	ProGrad	$60.42_{\pm 11.74}$	$64.22_{\pm 6.44}$	$68.56_{\pm 3.2}$	$74.65_{\pm 1.09}$	$74.93_{\pm 1.07}$
	BiomedCoOp (Ours)	$72.64_{\pm 2.41}$	$71.53_{\pm 1.5}$	$73.28_{\pm 2.30}$	$76.26_{\pm 0.38}$	$78.72_{\pm 0.23}$
CTKIDNEY	BiomedCLIP			42.43		
	BiomedCLIP + Ensemble			56.82		
	CLIP-Adapter	$47.17_{\pm 3.74}$	$41.94_{\pm 2.15}$	$42.19_{\pm 2.27}$	$44.64_{\pm 0.90}$	$47.28_{\pm 1.41}$
	Tip-Adapter	$45.85_{\pm 5.41}$	$51.65_{\pm 7.87}$	$55.33_{\pm 4.10}$	$69.89_{\pm 8.74}$	$73.38_{\pm 7.77}$
	Tip-Adapter-F	$46.68_{\pm 6.70}$	$58.99_{\pm 8.54}$	$60.18_{\pm 10.73}$	$75.24_{\pm 6.89}$	$82.07_{\pm 3.29}$
	Standard LP	$43.82_{\pm 6.43}$	$59.35_{\pm 6.49}$	$69.54_{\pm 7.67}$	$78.89_{\pm 7.37}$	$82.50_{\pm 5.22}$
	LP++	$57.70_{\pm 2.85}$	$61.57_{\pm 3.38}$	$65.73_{\pm 9.15}$	$77.06_{\pm 7.96}$	$79.07_{\pm 7.67}$
	CoOp	$54.51_{\pm 8.74}$	$60.57_{\pm 2.26}$	$68.12_{\pm 2.11}$	$77.4_{\pm 3.87}$	$83.52_{\pm 1.8}$
	CoCoOp	$47.88_{\pm 7.72}$	$52.71_{\pm 9.71}$	$61.07_{\pm 1.33}$	$73.93_{\pm 1.5}$	$77.7_{\pm 2.65}$
	KgCoOp	$58.92_{\pm 1.28}$	$62.81_{\pm 3.38}$	$68.68_{\pm 5.54}$	$77.43_{\pm 4.2}$	$77.67_{\pm 3.12}$
	ProGrad	$54.65_{\pm 8.97}$	$64.66_{\pm 5.31}$	$67.90_{\pm 2.02}$	$78.23_{\pm 4.74}$	$81.13_{\pm 2.28}$
	BiomedCoOp (Ours)	$56.13_{\pm 4.19}$	$64.21_{\pm 5.57}$	$66.5_{\pm 1.92}$	$77.16_{\pm 3.98}$	$83.20_{\pm 2.37}$
DermaMNIST	BiomedCLIP			38.75		
	BiomedCLIP + Ensemble			53.62		
	CLIP-Adapter	$35.96_{\pm 6.70}$	$36.01_{\pm 6.63}$	$34.97_{\pm 4.17}$	$34.28_{\pm 6.55}$	$29.02_{\pm 3.80}$
	Tip-Adapter	$37.52_{\pm 2.12}$	$40.98_{\pm 13.52}$	$47.31_{\pm 6.23}$	$61.67_{\pm 5.79}$	$62.67_{\pm 0.97}$
	Tip-Adapter-F	$37.34_{\pm 15.72}$	$38.52_{\pm 4.39}$	$50.44_{\pm 5.30}$	$43.87_{\pm 2.18}$	$53.86_{\pm 4.99}$
	Standard LP	$30.67_{\pm 13.12}$	$38.13_{\pm 10.28}$	$49.77_{\pm 8.34}$	$51.02_{\pm 2.99}$	$55.34_{\pm 3.56}$
	LP++	$26.93_{\pm 3.93}$	$26.16_{\pm 11.70}$	$36.29_{\pm 9.19}$	$45.78_{\pm 2.74}$	$50.98_{\pm 2.14}$
	CoOp	$25.88_{\pm 9.07}$	$38.92_{\pm 6.01}$	$43.71_{\pm 6.27}$	$46.8_{\pm 6.80}$	$51.07_{\pm 2.56}$
	CoCoOp	$24.51_{\pm 4.22}$	$24.96_{\pm 0.76}$	$25.29_{\pm 5.61}$	$40.42_{\pm 2.44}$	$40.97_{\pm 6.50}$
	KgCoOp	$27.1_{\pm 10.81}$	$30.28_{\pm 4.45}$	$35.35_{\pm 8.07}$	$38.79_{\pm 4.85}$	$36.59_{\pm 2.32}$
	ProGrad	$33.98_{\pm 10.76}$	$37.66_{\pm 6.74}$	$43.69_{\pm 10.96}$	$51.07_{\pm 2.47}$	$46.33_{\pm 5.13}$
	BiomedCoOp (Ours)	$58.64_{\pm 4.71}$	$57.17_{\pm 1.28}$	$60.07_{\pm 1.81}$	$61.98_{\pm 0.77}$	$62.59_{\pm 1.83}$

? S8? BiomedCoOp??????????????????????????(%)?

	BiomedCLIP			54.58		
	BiomedCLIP + Ensemble			57.5		
	CLIP-Adapter	$54.83_{\pm 0.48}$	$54.83_{\pm 0.48}$	$54.83_{\pm 0.48}$	$56.08_{\pm 0.86}$	$56.50_{\pm 1.00}$
	Tip-Adapter	$56.72_{\pm 3.42}$	$60.94_{\pm 5.30}$	$69.61_{\pm 2.06}$	$69.13_{\pm 1.44}$	$74.22_{\pm 1.51}$
	Tip-Adapter-F	$59.19_{\pm 0.89}$	$64.22_{\pm 3.24}$	$69.94_{\pm 2.28}$	$75.86_{\pm 1.00}$	$78.00_{\pm 1.06}$
Kvasir	Standard LP	$54.30_{\pm 2.04}$	$62.00_{\pm 0.81}$	$72.38_{\pm 2.65}$	$78.88_{\pm 0.73}$	$79.00_{\pm 0.81}$
	LP++	$58.27_{\pm 3.95}$	$60.47_{\pm 3.24}$	$69.36_{\pm 0.84}$	$72.52_{\pm 2.89}$	$75.41_{\pm 1.21}$
	CoOp	$58.2_{\pm 1.64}$	$64.86_{\pm 1.4}$	$70.78_{\pm 0.31}$	$77.14_{\pm 1.25}$	$77.88_{\pm 0.12}$
	CoCoOp	$59.45_{\pm 3.25}$	$65.5_{\pm 3.41}$	$68.94_{\pm 1.29}$	$72.92_{\pm 1.46}$	$75.22_{\pm 2.04}$
	KgCoOp	$61.67_{\pm 2.16}$	$65.67_{\pm 1.94}$	$68.28_{\pm 0.35}$	$72.05_{\pm 1.8}$	$72.95_{\pm 1.31}$
	ProGrad	$60.78_{\pm 0.24}$	$64.70_{\pm 0.53}$	$70.00_{\pm 0.24}$	$76.03_{\pm 1.50}$	$75.88_{\pm 0.95}$
	BiomedCoOp (Ours)	$62.17_{\pm 1.95}$	$67.25_{\pm 2.59}$	$74.08_{\pm 1.10}$	$77.72_{\pm 0.52}$	$78.89_{\pm 1.21}$
	BiomedCLIP			30.65		
	BiomedCLIP + Ensemble			31.52		
	CLIP-Adapter	$31.27_{\pm 0.69}$	$31.67_{\pm 0.88}$	$33.26_{\pm 0.39}$	$36.48_{\pm 1.32}$	$42.06_{\pm 2.40}$
	Tip-Adapter	$46.14_{\pm 9.62}$	$63.32_{\pm 2.58}$	$70.05_{\pm 1.11}$	$69.57_{\pm 1.63}$	$77.68_{\pm 1.42}$
	Tip-Adapter-F	$52.81_{\pm 3.10}$	$58.90_{\pm 4.95}$	$71.74_{\pm 2.72}$	$74.51_{\pm 2.43}$	$80.43_{\pm 2.85}$
CHMNIST	Standard LP	$58.44_{\pm 2.02}$	$64.42_{\pm 3.81}$	$71.07_{\pm 2.23}$	$76.30_{\pm 3.22}$	$80.34_{\pm 1.83}$
	LP++	$57.18_{\pm 6.46}$	$60.61_{\pm 1.26}$	$67.79_{\pm 6.97}$	$72.40_{\pm 0.85}$	$78.32_{\pm 1.48}$
	CoOp	$57.34_{\pm 4.2}$	$59.68_{\pm 1.12}$	$68.66_{\pm 2.14}$	$75.00_{\pm 0.82}$	$79.63_{\pm 1.26}$
	CoCoOp	$49.07_{\pm 4.41}$	$50.82_{\pm 3.41}$	$58.58_{\pm 2.15}$	$66.58_{\pm 1.14}$	$72.16_{\pm 0.52}$
	KgCoOp	$59.02_{\pm 4.1}$	$60.06_{\pm 1.12}$	$68.77_{\pm 1.02}$	$69.50_{\pm 0.07}$	$73.58_{\pm 1.19}$
	ProGrad	$60.15_{\pm 5.76}$	$59.60_{\pm 1.53}$	$69.13_{\pm 1.39}$	$70.99_{\pm 0.36}$	$75.11_{\pm 1.50}$
	BiomedCoOp (Ours)	$59.82_{\pm 2.43}$	$59.79_{\pm 1.36}$	$71.19_{\pm 1.74}$	$74.78_{\pm 1.19}$	$79.05_{\pm 2.24}$
	BiomedCLIP			50.03		
	BiomedCLIP + Ensemble			61.84		
	CLIP-Adapter	$54.83_{\pm 0.36}$	$53.47_{\pm 2.95}$	$52.91_{\pm 1.70}$	$56.33_{\pm 0.45}$	$57.56_{\pm 1.13}$
	Tip-Adapter	$75.37_{\pm 4.02}$	$72.73_{\pm 8.09}$	$83.32_{\pm 3.95}$	$87.25_{\pm 1.75}$	$89.17_{\pm 0.41}$
	Tip-Adapter-F	$74.21_{\pm 4.35}$	$71.82_{\pm 7.31}$	$79.57_{\pm 10.02}$	$90.41_{\pm 2.43}$	$92.35_{\pm 1.08}$
	Standard LP	$74.50_{\pm 2.61}$	$78.40_{\pm 7.36}$	$85.30_{\pm 3.56}$	$90.24_{\pm 0.41}$	$92.77_{\pm 1.17}$
	LP++	$63.05_{\pm 9.52}$	$71.42_{\pm 3.04}$	$82.61_{\pm 2.31}$	$89.14_{\pm 2.07}$	$92.58_{\pm 0.38}$
	CoOp	$71.90_{\pm 3.53}$	$76.55_{\pm 2.81}$	$84.66_{\pm 2.26}$	$87.50_{\pm 0.26}$	$92.19_{\pm 0.48}$
	CoCoOp	$63.66_{\pm 4.49}$	$71.76_{\pm 0.55}$	$77.44_{\pm 2.47}$	$85.57_{\pm 1.83}$	$87.38_{\pm 0.52}$
	KgCoOp	$71.80_{\pm 2.13}$	$75.18_{\pm 1.05}$	$82.10_{\pm 2.35}$	$84.63_{\pm 0.30}$	$86.79_{\pm 0.53}$
	ProGrad	$72.48_{\pm 3.22}$	$74.76_{\pm 1.40}$	$84.72_{\pm 2.85}$	$87.86_{\pm 0.70}$	$90.70_{\pm 0.66}$
	BiomedCoOp (Ours)	$77.56_{\pm 2.84}$	$77.74_{\pm 2.00}$	$85.60_{\pm 1.61}$	$88.77_{\pm 1.14}$	$92.68_{\pm 0.57}$
	BiomedCLIP			26.26		
	BiomedCLIP + Ensemble			39.27		
	CLIP-Adapter	$25.49_{\pm 0.46}$	$25.49_{\pm 0.46}$	$26.07_{\pm 0.46}$	$25.84_{\pm 0.87}$	$26.05_{\pm 0.43}$
	Tip-Adapter	$26.52_{\pm 0.42}$	$31.07_{\pm 3.84}$	$43.42_{\pm 7.04}$	$48.08_{\pm 7.40}$	$54.23_{\pm 5.13}$
	Tip-Adapter-F	$39.53_{\pm 10.83}$	$33.07_{\pm 5.63}$	$47.37_{\pm 6.70}$	$56.07_{\pm 2.57}$	$62.85_{\pm 1.10}$
RETINA	Standard LP	$39.35_{\pm 6.96}$	$46.03_{\pm 0.79}$	$51.31_{\pm 6.52}$	$53.94_{\pm 1.98}$	$62.27_{\pm 2.80}$
	LP++	$35.77_{\pm 5.75}$	$39.37_{\pm 7.35}$	$46.95_{\pm 10.07}$	$53.44_{\pm 1.95}$	$60.62_{\pm 1.46}$
	CoOp	$35.02_{\pm 1.40}$	$35.26_{\pm 3.34}$	$42.22_{\pm 3.09}$	$51.87_{\pm 1.78}$	$59.38_{\pm 0.87}$
	CoCoOp	$32.94_{\pm 0.75}$	$36.43_{\pm 4.05}$	$39.75_{\pm 3.99}$	$48.45_{\pm 1.39}$	$53.91_{\pm 1.52}$
	KgCoOp	$33.54_{\pm 2.77}$	$35.17_{\pm 2.48}$	$42.61_{\pm 3.16}$	$49.97_{\pm 2.24}$	$51.18_{\pm 1.66}$
	ProGrad	$33.49_{\pm 1.98}$	$36.49_{\pm 4.64}$	$43.09_{\pm 3.89}$	$52.26_{\pm 2.38}$	$50.47_{\pm 2.40}$
	BiomedCoOp (Ours)	$36.64_{\pm 3.34}$	$38.67_{\pm 1.79}$	$45.58_{\pm 5.03}$	$56.47_{\pm 1.37}$	$61.28_{\pm 1.06}$
	BiomedCLIP			29.53		
	BiomedCLIP + Ensemble			39.37		
	CLIP-Adapter	$29.00_{\pm 0.17}$	$28.66_{\pm 0.45}$	$28.96_{\pm 0.46}$	$28.80_{\pm 0.20}$	$29.08_{\pm 0.32}$
	Tip-Adapter	$29.04_{\pm 0.67}$	$33.55_{\pm 5.96}$	$24.19_{\pm 4.23}$	$25.76_{\pm 3.35}$	$33.17_{\pm 7.59}$
	Tip-Adapter-F	$30.01_{\pm 0.50}$	$28.38_{\pm 2.18}$	$26.59_{\pm 5.70}$	$26.46_{\pm 2.20}$	$27.67_{\pm 3.21}$
KneeXray	Standard LP	$26.02_{\pm 11.08}$	$26.57_{\pm 5.17}$	$27.83_{\pm 4.92}$	$22.20_{\pm 3.68}$	$23.97_{\pm 3.55}$
	LP++	$21.25_{\pm 8.60}$	$26.40_{\pm 3.26}$	$28.92_{\pm 4.97}$	$23.75_{\pm 2.50}$	$26.38_{\pm 3.39}$
	CoOp	$24.96_{\pm 9.41}$	$25.89_{\pm 5.06}$	$23.85_{\pm 4.25}$	$26.23_{\pm 4.01}$	$28.48_{\pm 1.84}$
	CoCoOp	$25.42_{\pm 6.38}$	$28.85_{\pm 8.24}$	$30.66_{\pm 4.49}$	$21.78_{\pm 8.29}$	$24.86_{\pm 4.15}$
	KgCoOp	$29.07_{\pm 3.31}$	$28.14_{\pm 4.53}$	$22.44_{\pm 2.88}$	$23.37_{\pm 3.35}$	$24.80_{\pm 0.47}$
	ProGrad	$30.09_{\pm 6.00}$	$23.83_{\pm 0.57}$	$23.95_{\pm 2.87}$	$24.78_{\pm 2.32}$	$26.27_{\pm 3.29}$
	BiomedCoOp (Ours)	$36.13_{\pm 1.75}$	$37.72_{\pm 0.54}$	$35.91_{\pm 0.54}$	$37.7_{\pm 1.00}$	$39.69_{\pm 1.75}$

? S8????BiomedCoOp??????????????????????????????????%??

OCTMNIST	BiomedCLIP			30.00		
	BiomedCLIP + Ensemble			47.40		
	CLIP-Adapter	$44.00_{\pm 5.79}$	$49.73_{\pm 2.41}$	$49.96_{\pm 1.77}$	$49.50_{\pm 3.33}$	$52.73_{\pm 0.62}$
	Tip-Adapter	$32.36_{\pm 3.94}$	$33.8_{\pm 6.16}$	$38.10_{\pm 5.01}$	$53.93_{\pm 3.17}$	$53.33_{\pm 3.92}$
	Tip-Adapter-F	$46.66_{\pm 2.58}$	$53.93_{\pm 1.67}$	$55.20_{\pm 4.75}$	$65.00_{\pm 6.61}$	$72.50_{\pm 1.38}$
	Standard LP	$47.25_{\pm 12.64}$	$54.21_{\pm 8.23}$	$61.00_{\pm 7.07}$	$65.85_{\pm 9.01}$	$69.40_{\pm 3.68}$
	LP++	$47.24_{\pm 13.84}$	$53.18_{\pm 9.08}$	$59.02_{\pm 8.59}$	$63.69_{\pm 8.26}$	$68.35_{\pm 7.42}$
	CoOp	$52.63_{\pm 2.95}$	$53.57_{\pm 3.86}$	$53.37_{\pm 2.35}$	$63.67_{\pm 4.47}$	$65.47_{\pm 7.47}$
	CoCoOp	$49.33_{\pm 4.58}$	$50.93_{\pm 8.01}$	$48.57_{\pm 6.25}$	$55.40_{\pm 1.88}$	$60.67_{\pm 3.41}$
	KgCoOp	$50.63_{\pm 3.18}$	$50.53_{\pm 5.39}$	$52.97_{\pm 1.58}$	$61.03_{\pm 3.78}$	$62.80_{\pm 3.85}$
	ProGrad	$51.40_{\pm 3.05}$	$55.33_{\pm 3.38}$	$55.07_{\pm 1.22}$	$62.17_{\pm 6.01}$	$63.33_{\pm 6.15}$
	BiomedCoOp (Ours)	$51.83_{\pm 1.52}$	$55.03_{\pm 4.72}$	$54.73_{\pm 1.86}$	$58.87_{\pm 5.35}$	$66.93_{\pm 2.13}$
Average	BiomedCLIP			42.05		
	BiomedCLIP + Ensemble			52.27		
	CLIP-Adapter	$44.66_{\pm 2.97}$	$43.91_{\pm 2.48}$	$44.36_{\pm 1.94}$	$45.42_{\pm 2.38}$	$46.69_{\pm 1.71}$
	Tip-Adapter	$49.19_{\pm 4.84}$	$52.36_{\pm 6.57}$	$57.33_{\pm 5.07}$	$61.98_{\pm 5.76}$	$67.15_{\pm 4.25}$
	Tip-Adapter-F	$51.17_{\pm 8.33}$	$52.74_{\pm 5.88}$	$61.23_{\pm 6.22}$	$65.91_{\pm 3.64}$	$70.91_{\pm 2.65}$
	Standard LP	$47.25_{\pm 8.65}$	$54.21_{\pm 7.80}$	$61.00_{\pm 6.81}$	$65.85_{\pm 4.89}$	$69.40_{\pm 2.91}$
	LP++	$47.24_{\pm 7.68}$	$53.18_{\pm 7.29}$	$59.02_{\pm 6.93}$	$63.69_{\pm 4.68}$	$68.35_{\pm 3.59}$
	CoOp	$50.16_{\pm 6.93}$	$54.18_{\pm 4.31}$	$59.75_{\pm 3.72}$	$65.84_{\pm 3.66}$	$69.62_{\pm 2.83}$
	CoCoOp	$48.49_{\pm 4.39}$	$51.28_{\pm 5.06}$	$54.69_{\pm 4.79}$	$61.08_{\pm 3.49}$	$65.09_{\pm 2.87}$
	KgCoOp	$51.83_{\pm 5.53}$	$53.47_{\pm 5.07}$	$58.59_{\pm 4.50}$	$63.65_{\pm 2.73}$	$64.88_{\pm 1.95}$
	ProGrad	$51.88_{\pm 6.39}$	$54.71_{\pm 4.46}$	$60.42_{\pm 4.78}$	$65.61_{\pm 3.02}$	$67.13_{\pm 3.00}$
	BiomedCoOp (Ours)	$57.03_{\pm 2.80}$	$59.13_{\pm 3.64}$	$63.95_{\pm 2.42}$	$68.32_{\pm 2.65}$	$72.42_{\pm 1.62}$

? S8???BiomedCoOp??????????????????????????????????%??

Dataset	Context Token #1	Context Token #2	Context Token #3	Context Token #4
BTMRI	`mri` (2.4971)	`curcumin` (2.5835)	`of` (1.5667)	`a` (1.6353)
BUSI	`a` (2.5550)	`photo` (3.5649)	`of` (2.1298)	`b` (3.4897)
COVID-QU-Ex	`measured` (2.1999)	`image` (2.2856)	`of` (1.9166)	`a` (1.9205)
CTKIDNEY	`a` (2.1290)	`schem` (2.6564)	`right` (2.3790)	`a` (1.7574)
DermaMNIST	`dextrose` (2.8292)	`photo` (3.1084)	`ricin` (3.2378)	`autologous` (3.0297)
Kvasir	`endoscopy` (2.1880)	`scar` (2.4835)	`of` (2.2698)	`maintained` (2.4771)
CHMNIST	`a` (3.0301)	`original` (3.4248)	`composed` (2.2125)	`discern` (3.4506)
LC25000	`a` (1.5298)	`photo` (2.3540)	`of` (1.6363)	`a` (2.0292)
RETINA	`a` (1.5986)	`papill` (2.3636)	`of` (1.6976)	`receptive` (2.1135)
KneeXray	`a` (4.2063)	`calcification` (5.4999)	`osteoc` (2.8673)	`showed` (2.9774)
OCTMNIST	`localized` (2.1744)	`example` (3.6750)	`of` (1.8752)	`possible` (2.4803)

?S9? BiomedCoOp???44???