Pattern-Anchored Adaptive Prototype Learning for Gastroscopic Lesion Detection and Beyond

Xuanye Zhang[1†], Xiaoqing Hu[2†], Guanbin Li[3], Si-Qi Liu[1‡], Xiang Wan[1], and Yuanhuan Xiong[2‡]

[1] Shenzhen Research Institute of Big Data, CUHK-Shenzhen, China
[2] Jiangxi Provincial Maternal and Child Health Hospital, Nanchang, Jiangxi, China
[3] School of Computer Science and Engineering, Sun Yat-sen University, China

Abstract. Gastroscopic Lesion Detection (GLD) is one of the critical tasks within computer-assisted gastroscopic diagnostics. Endoscopists adopt a pattern-based philosophy for GLD: they identify and summarize typical sub-category patterns with specific medical meanings and conduct GLD based on these patterns. However, the current gastroscopic lesion detectors follow the classical data-driven deep-learning-based training paradigm, which differs from the endoscopists' diagnosis process and leads to low interpretability, limiting their performance and potential for daily clinical practice and patient care. The intuitive data-driven solution with sub-category pattern labels may work but it requires expensive annotation costs. In this work, we imitate the pattern-based philosophy with limited labels and propose a Pattern-Anchored Adaptive Prototype Learning (PAAPL) for Gastroscopic Lesion Detection. PAAPL consists of a Prototype-based Gastroscopic Lesion Detector (PGLD) and a Pattern-Anchored Adaptive Learning (PAAL) strategy. PGLD achieves sub-category pattern detection based on similarity to prototypes. PAAL proposes a vector-wise prototype formulation and an adaptive prototype update strategy to anchor prototypes to limited-annotated patterns with specific medical meanings and adaptively learn pattern characteristics from unannotated data in GLD datasets. We evaluate PAAPL on the LGLDD and Endo21 datasets, demonstrating its ability to learn and detect sub-category patterns trained with limited annotations. By doing this, PAAPL enhances detector interpretability and yields significant performance improvement (+3.7AP on LGLDD/+5.4AP on Endo21).

Keywords: Gastroscopic Lesion Detection · Sub-category Patterns · Limited-annotated data

1 Introduction

Gastroscopy plays a crucial role in daily clinical practice for diagnosing a spectrum of gastropathy, including early cancer detection [16], gastric polyp detec-

†The two authors contribute equally to this paper.
‡Corresponding authors: Si-Qi Liu (siqiliu@sribd.cn) and Yuanhuan Xiong (xiongyuanhuan@163.com)

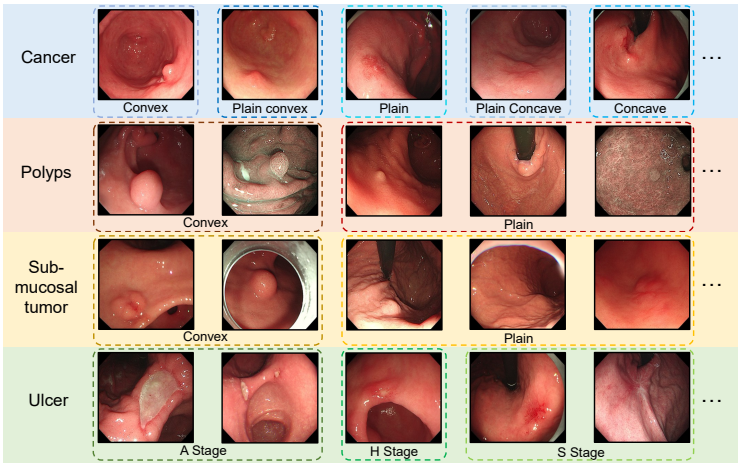

Fig. 1: **Typical patterns for each category of Gastroscopic Lesions:** Each category of gastroscopic lesions has multiple typical sub-category patterns and appears large intra-category morphological diversity of convexity, concavity, or texture.

tion [1], gastric varices detection [6], etc. Gastroscopic Lesion Detection (GLD) is a key task within this domain, which aims to locate and identify different gastroscopic lesions with cheaper bounding box annotations than segmentation. Compared with colonoscopic lesions, gastroscopic lesions exhibit higher sub-category diversity, particularly in terms of convexity and concavity [5, 4](see Fig. 1). The classical data-driven deep-learning-based training paradigm adopted by the current gastroscopic lesion detectors [10, 8, 11, 3, 15, 16, 7, 13] makes them pay less attention to these diverse sub-category lesion patterns, which differs from endoscopists' behaviors. Due to this reason, the current gastroscopic lesion detectors may suffer from low interpretability, limiting their potential for further improvement and application in daily clinical practice and patient care.

We examine how endoscopists learn to conduct GLD and discover a more interpretable pattern-based philosophy [5, 4]: they identify and summarize typical sub-category patterns with specific medical meanings for each category (see Fig. 1). With this knowledge, they can handle pronounced sub-category morphological diversity with limited prototype examples from textbooks and refine their diagnostic precision with more examples in clinical practice. The most intuitive way to integrate such philosophy into GLD is re-annotating the current GLD datasets with sub-category pattern labels and re-training the detectors. However, annotating exhaustive sub-category labels (well-annotated data) is extremely expensive. Few-shot learning [9, 14, 12] could be a cheaper solution but it may not be suitable in the GLD setting because the limited-annotated sub-category pattern labels are not mutually exclusive with well-annotated category labels.

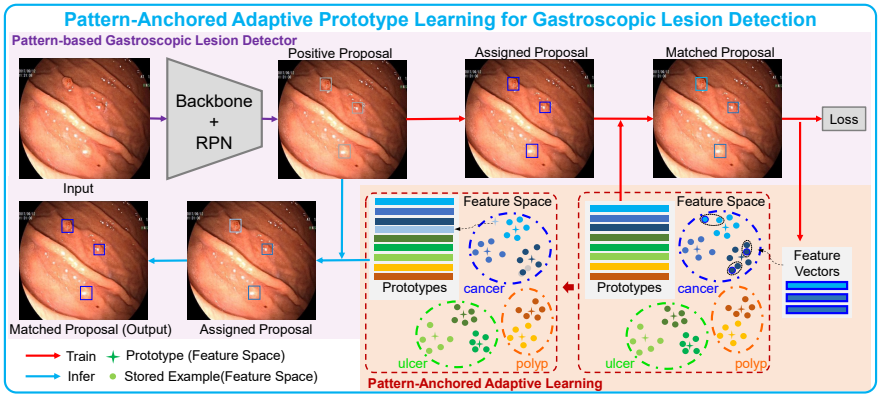

Fig. 2: **Pipeline of Pattern-Anchored Adaptive Prototype Learning (PAAPL) for GLD:** PAAPL consists of a Prototype-based Gastroscopic Lesion Detector (PGLD) and Pattern-Anchored Adaptive Learning (PAAL) strategy. (→: Training process, →: Inferring process)

In this work, we propose Pattern-Anchored Adaptive Prototype Learning (PAAPL) for GLD that can integrate the pattern-based philosophy with limited sub-category pattern annotations (non-mutually-exclusive to category labels). PAAPL contains two main components: Prototype-based Gastroscopic Lesion Detector (PGLD) to achieve prototype-based sub-category pattern detection and Pattern-Anchored Adaptive Learning (PAAL) strategy to learn sub-category pattern prototypes from GLD datasets with limited annotations. Specifically, PGLD introduces a Prototype Branch following the Region Proposal Network (RPN) for prototype-based pattern detection. This brunch employs contrastive clustering to learn prototypes of sub-category patterns and achieves detection based on the similarity to the pattern prototypes. The PAAL formulates prototypes using feature vectors of both limited-annotated patterns and representative examples in GLD datasets to anchor the prototypes with specific patterns and learn more representative characteristics of patterns. Moreover, PAAL adaptively updates prototypes based on different similarity threshold to preserve stable updates and avoid potential forgetting of some patterns.

We evaluate the effectiveness of PAAPL on the LGLDD and Endo21 datasets. Experimental results demonstrate that PAAPL can learn prototypes anchoring to specific sub-category gastroscopic lesion patterns from GLD datasets with limited annotations, which improves the detector interpretability and yields significant improvement (LGLDD: +3.7AP/Endo21: +5.4AP).

2 Methodology

The proposed Pattern-Anchored Adaptive Prototype Learning (PAAPL) for GLD consists of a Prototype-based Gastroscopic Lesion Detector (PGLD, Sec.2.1)

and a Pattern-Anchored Adaptive Learning (PAAL) strategy. The PAAL adopts a novel prototype formulation (Sec. 2.2) and update (Sec. 2.3) strategy. Fig. 2 demonstrates the pipeline of PAAPL.

2.1 Prototype-based Gastroscopic Lesion Detector

The PGLD is designed for prototype-based pattern-level detection. PGLD adopts the classical 2-stage detector framework of Faster RCNN [10] and introduces a novel Prototype Branch after Region Proposal Network (RPN) to achieve the goal. In the Prototype Branch, PGLD maintains a set of prototypes denoted as $\mathcal{P} = \{p_{11}, p_{12}, ..., p_{c1}, p_{c2}, ...\}$ to facilitate pattern-level detection, where $c \in C$ and C represents the number of categories in the datasets. The size of prototypes in $\mathcal{P}$ and the formulation of prototypes can both be flexible.

Given an input gastroscopic lesion images I, PGLD first extracts positive proposals $\mathcal{O} = \{o_1, o_2, ...\}$, from the Region Proposal Network (RPN), akin to Faster RCNN. In the following, the training and inferring process are different.

Training process (→ in Fig. 2): 1) For each o_i (its feature vector denotes as f_i), PGLD assigns the ground truth to it, obtaining assigned proposals o_{ia} and assigned label c_i. 2) For each assigned proposal o_{ia}, PGLD obtains prototypes $\mathcal{P}_c$ of c_i, where $\mathcal{P}_c = \{p_{ij} | i = c_i\}$, computes the similarity $s_{cj} = sim(f_i, p)$ for $p \in \mathcal{P}_c$, and obtain similarity sets: $\mathcal{S}_i = \{s_{c1}, s_{c2}, ...\}$, 3) PGLD matches the most similar prototype for $p \in \mathcal{P}_c$, and obtain p_i and matched proposal o_{im}, which can be expressed as: $p_i = \{p_{ij} | j = \arg\max_{s \in \mathcal{S}_i} s_{ij}\}$, where the similarity function $sim(\cdot, \cdot)$ is: $sim(x, y) = \frac{x \cdot y}{\|x\| \|y\|}$.

For each matched proposal o_{im}, PGLD sends them to Classification Brunch, Box Regression Brunch, and Prototype Brunch to calculate the loss. The f_i is also used to update the prototype p_i using some rules. We employ the InfoNCE Loss for effective contrastive learning to Prototype Brunch $\mathcal{L}_p$:

$$\mathcal{L}_p = -log \frac{exp(f_i \cdot p_i / \tau)}{\sum_{p_{ij} \in \mathcal{P}} exp(f_i \cdot p_{ij} / \tau)},$$

where τ is a temperature hyper-parameter and is set to 0.07.

Inferring process (→ in Fig. 2): 1) For each o_i, PGLD computes the $s_{cj} = sim(f_i, p)$ for $p \in \mathcal{P}$ and obtain $\mathcal{S}_i = \{s_{11}, s_{12}, ...\}$. 2) PGLD first matches the most similar prototype to each positive proposal $c, j = \arg\max_{s \in \mathcal{S}_i} s_{cj}$, and obtains matched proposal o_{im}. 3) For each o_{im}, PGLD assigns the labels of most similar prototypes to it ($\hat{c}_i = c$) and obtains assigned proposals o_{ia}. 4) PGLD regresses the bounding box for o_{ia} and obtains detection outputs.

Though the Prototype Branch can predict labels and assign ground truth, the classification branch remains essential for PGLD, particularly in the early stages of training when the pattern quality may be suboptimal, as the classification branch helps stabilize the training process.

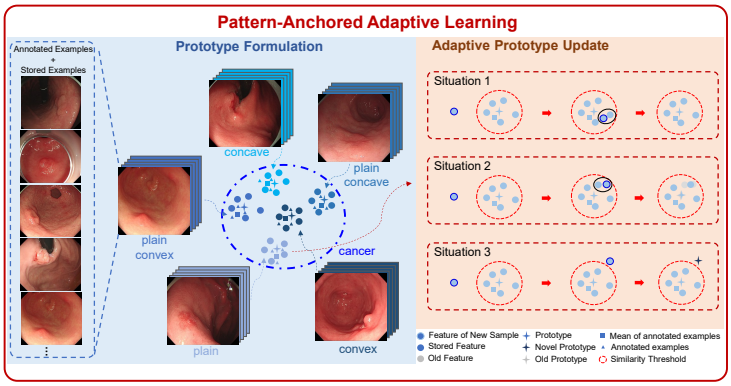

Fig. 3: **Pipeline of Pattern-Anchored Adaptive Learning (PAAL):** PAAL formulates prototypes with both limited-annotated examples and data examples of GLD datasets. The prototype update of PAAL can be categorized into 3 situations.

2.2 Prototype Formulation of PAAL

The PAAL requires prototypes anchoring to specific patterns and being able to represent examples of GLD datasets.

Formulation: PAAL formulates pattern prototypes as follows: 1) PAAL stores feature vectors of examples with pattern-level annotations (denoted as $\mathcal{A} = \{a_1, a_2, ...\}$) to anchor the pattern. The mean of the $\mathcal{A}$ is denoted as a_m. 2) PAAL maintains a repository of up-to-N feature vectors of examples (denoted as $\mathcal{E} = \{e_1, e_2, ..., e_N\}$) from the GLD dataset for each pattern, and the mean of $\mathcal{E}$ is denoted as e_m.

Initialization: PAAL uses the $\mathcal{A}$ to initialize $\mathcal{E}$ of each pattern. N is larger than the size of $\mathcal{A}$. Once $\mathcal{E}$ is obtained, PAAL initializes the pattern prototype using the mean e_m of this unfilled $\mathcal{E}$ as p_i. In the learning process, it directly fills the feature vector to $\mathcal{E}$ until the size of $\mathcal{E}$ is equal to N.

Training: Except for the contrastive learning between f_i and p_i in Sec. 2.1, PAAL also computes the similarity loss between f_i and corresponding a_{mi} to anchor the prototype by anchoring the examples to obtain $\mathcal{L}_a$, where $\mathcal{L}_a = sim(f_i, a_{mi})$). With $\mathcal{L}_a$ incorporated, the overall loss $\mathcal{L}_d$ for prototype brunch becomes $\mathcal{L}_d = \mathcal{L}_p + \mathcal{L}_a$ The total Loss function can be expressed as:

$$\mathcal{L} = \mathcal{L}_{PRN} + \phi_d \mathcal{L}_d + \phi_{box} \mathcal{L}_{box} + \phi_{cls} \mathcal{L}_{cls}$$

where $\mathcal{L}_{PRN}$, $\mathcal{L}_{box}$, $\mathcal{L}_{cls}$ are the loss of RPN, box regression brunch and classification brunch, respectively. ϕ_d, ϕ_{box}, ϕ_{cls} are the loss weight of $\mathcal{L}_d$, $\mathcal{L}_{box}$, $\mathcal{L}_{cls}$

This formulation ensures that the learned pattern prototypes anchor to specific patterns and learning the characteristics of data examples in GLD datasets.

2.3 Prototype Update of PAAL

The prototype update of PAAL need to balance preserving update stability and avoiding potential omitting or forgetting for some examples.

In the beginning, given a feature vector f_i of a lesion, PAAL measures the similarity $sim(f_i, p_i)$ against a predefined threshold τ_d:

If $sim(f_i, p_i) > \tau_d$, it indicates a high similarity to existing prototypes. PAAL then evaluates the similarity between f_i and $e_i \in \mathcal{E}$ to identify the most similar example e_s, where $e_s = \underset{e_i \in \mathcal{E}}{max}\, sim(f_i, e_i)$, Subsequently, PAAL compares f_i, e_s with prototype p_i and identifies the less similar between them as the more unique example e_u, where $e_u = \underset{f \in \{f_i, e_s\}}{arg\ max}\, sim(f, p_i)$,

Situation 1: $e_u = e_i$, PAAL remains the e_s unchanged ($e_s = e_i$).
Situation 2: $e_u = f_i$, PAAL employs a momentum approach to update the stored examples: $e_s = m \cdot e_u + (1-m) \cdot e_s$, where m is the momentum coefficient.
Situation 3: $sim(f_i, p_i) < \tau_d$, which indicates a limited similarity to existing prototypes, PAAL adaptively splits the lesion as a discovered pattern from raw pattern, generates prototype for it, but keeps the same pattern label with raw pattern. By doing this, PAAL mitigates the risks of overlooking or forgetting examples and reduces the instability of prototypes. Specifically, PAAL regards f_i as the prototype p_d for this discovered pattern, incorporates p_d into the set of pattern prototypes $\mathcal{P}$, and preserves the raw medical meaning by inheriting the set of $\mathcal{A}_i$. This discovered pattern is updated within the framework above. The prototype update strategy can preserve stability for different initialization.

3 Experiment

Settings: Backbone: ResNet-50, Optimizer: Adam, Learning Rate: 1e-4, $\tau_d = 0.4$, $N = 20$. Training schedules, Faster-RCNN Loss, other settings follow SSL [16].

Datasets: We evaluate the PAAPL on LGLDD [16] and Endo21 [2]. **LGLDD** [16] comprises 12,292 lesion boxes, 10,083 gastroscopic images, and includes four types of lesions: polyp ('pol'), ulcer ('uls'), cancer ('can'), and sub-mucosal tumor ('stm'). **Endo21** [2] Sub-task 2 contains 1473 images and contain 'polyp'.

Sub-category Gastroscopic Lesion Pattern Annotations: Experienced endoscopists help us to identify typical sub-category patterns from morphological aspects. As is shown in Fig. 1, 1) cancers contain 5 typical sub-category patterns (concave, plain concave, plain, plain convex, and convex); 2) polyps contain 2 typical sub-category patterns (plain and convex); 3) ulcers contain 3 sub-category patterns (A stage, H stage, and S stage); and 4) smts contain 2 sub-category patterns (plain and convex). Doctors give us at least 7 typical examples for each pattern. Finally, we use 84 annotated examples in total for PAAPL.

Main Results: We report Quantitative (Tab. 1/Tab. 2(a)), t-SNE (Fig. 4.'PAAL'), and Qualitative (Fig. 5) results: 1) PAAPL can learn all the sub-category patterns (Fig. 4.PAAL) and can detect challenging pattern lesions (Fig. 5). Detecting these pattern lesions brings significant improvements (AP & Category-wise

Table 1: **Quantitative Results on LGLDD datasets:** PAAPL can bring significant performance enhancement (+3.7AP) compared to Faster-RCNN and overperform other methods, especially for the most challenging label 'can'.

Method	AP	AP50	AP75	pol	smt	uls	can
CenterNet [3]	29.3	57.2	25.4	41.6	36.0	27.3	12.1
Faster RCNN [10]	34.1	70.6	28.1	44.0	44.4	24.0	24.2
DINO [15]	32.2	66.6	26.6	42.0	43.8	22.2	20.6
HSL [16]	36.4	74.0	31.4	43.7	48.0	26.1	27.6
SSL [16]	37.3	74.8	33.2	44.9	51.0	26.1	27.3
STFT [13]	36.2	74.1	31.2	45.4	53.4	23.1	22.8
ECC [7]	36.7	74.1	32.8	**45.7**	**53.8**	23.3	24.1
Faster RCNN + FSFT	33.7 (-0.4)	69.8	27.6	43.6	44.2	23.2	23.7
PGLD	36.5 (+2.4)	73.6	32.7	44.2	49.1	25.7	27.5
PAAPL	**37.8 (+3.7)**	**75.7**	**33.6**	45.1	50.2	**26.6**	**29.4**

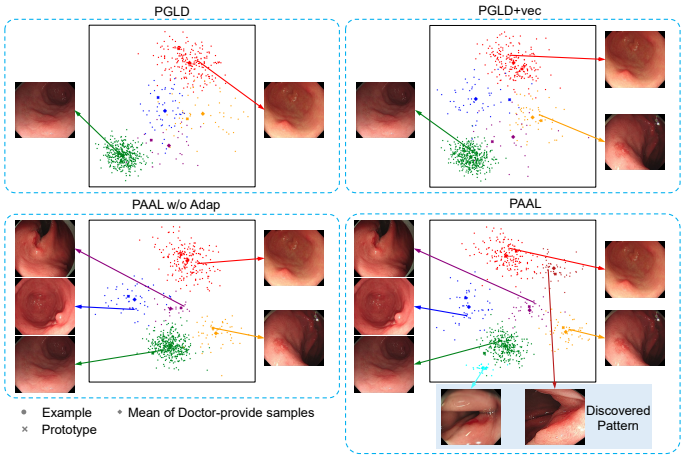

Fig. 4: **tSNE Results of label 'can' for different PAAL Design.**

AP) compared to Faster RCNN [10] (+3.7AP on LGLDD/+5.4AP on Endo21) and outperform some other classical general detector like CenterNet [3] and DINO [15] (Tab. 1 and Tab. 2(a)), 2) **Category-wise AP** also proves the AP enhancement source of PAAPL is improving detecting challenging category 'can' and 'uls' ('can' and 'uls' both have more sub-category patterns, including 'concave' sub-category patterns.), 3) **Limited Sub-category Annotations** help PAAPL outperform SSL [16], which uses large-scale unlabeled data and huge GPU time for backbone pre-training and semi-supervised learning, 4) **Comparison with few-shot detectors:** We train 2-stage few-shot detector [9] using category labels as prior knowledge and sub-category annotations for few-shot learning. Experiment results ('Faster-RCNN+FSFT' in Tab. 1) prove unsuitable usage of limited sub-category pattern annotations brings negative effects.

Comparison with Coloniscopic Polyp Detector: Compared to SOTA Coloniscopic Polyp Detector (CPD) (ECC [7] and STSF [13]), PAAPL can outperform

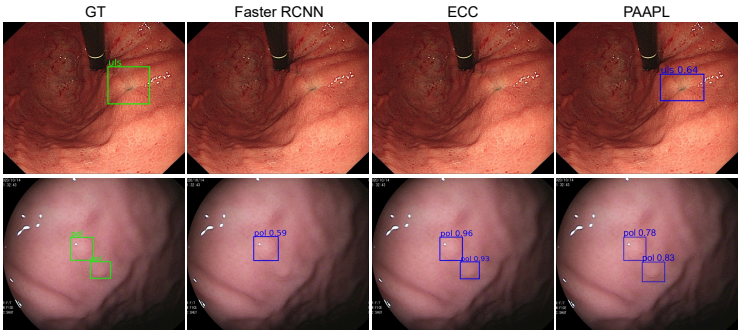

Fig. 5: **Qualitative Result:** PAAPL can detect concave pattern lesions and achieve comparable performance to ECC [7] in detection convex lesions.

Table 2: **Quantitative Results of:** (a) Endo21 datasets (b) Ablation Study for PAAL Design and Detector Interpretability on LGLDD datasets.

(a)

	AP	AP_{50}	AP_{75}
DETR	55.3	74.1	64.3
Faster RCNN	57.8	79.1	68.1
DINO	59.4	79.8	68.7
YOLO v5	60.5	81.0	66.4
SSL	61.9	83.0	69.2
PAAPL	**63.2 (+5.4)**	**84.4**	**69.7**

(b)

Method	AP	AP50	AP75
PGLD	36.5	73.6	32.7
PGLD+vec	36.8	74.1	32.8
PAAL w/o Adap	36.9	74.2	32.8
PAAL	**37.4**	**75.1**	**33.1**

them on mAP (Tab. 1). From category-wise AP, CPDs perform well on labels 'pol' and 'smt' but below expectations on labels 'can' and 'ulc'. This is because the typical sub-category patterns of 'pol' and 'smt' are 'convex' or 'plain', which is quite similar to colonoscopic polys and CPD aims to improve convex-like lesion detection. However, PAAPL demonstrates stronger adaptability in detecting different subcategory lesion patterns. Fig. 5 also reach similar conclusion.

Ablation Study for PAAL Design and Detector Interpretability: Experiments to evaluate designs of PAAL include: 1) 'PGLD': w/o any sub-category annotations, 2) 'PGLD+vec': PGLD + vector-wise prototype formulation (w/o anchoring to sub-category patterns), 3) 'PAAL w/o Adap': PAAL (anchoring to sub-category patterns) but w/o adaptive prototype update, and 4) 'PAAL'. The more discriminative pattern clusters learned by PAAL prove its performance enhancement sources from recognizing these patterns, improving interpretability.

According to Tab. 2(b) and t-SNE results of challenging label 'can' (Fig. 4), 1) vector-wise prototype can better learn and represent sub-category pattern ('PGLD+vec' learns 1 more sub-category pattern and enhance 0.3AP compared to 'PGLD'); 2) Though anchoring prototypes can help learn all the sub-category patterns and enhance model interpretability, adaptive prototype update strategy can further enhance performance ('PAAL' enhance 0.5AP to 'PAAL w/o Adap').

Conclusion: By integrating the endoscopists' pattern-based philosophy, PAAPL can learn to detect sub-category lesion patterns with affordable annotations, enhancing interpretability and overperform somemore expensive detector. Moreover, the ResNet-50 backbone ensure PAAPL can be deployed to cheaper device, which is significant to ground the model to some hospitals in rural areas. Finally, PAAPL demonstrates extensive potential and value of integrating valuble endoscopists' experience into deep-learning model, and provides affordable solutions to satisfy the increasing demands of detecting more sub-category lesion patterns in clinical practice (annotating limited examples and training with PAAPL).

Acknowledgments. This work is supported by the Guangdong Basic and Applied Basic Research Foundation (Grant Nos. 2025A1515011556 and 2023A1515110706), the Project (No. 20232ABC03A25), the Shenzhen Science and Technology Program (JCYJ20220818103001002), the Guangdong Provincial Key Laboratory of Big Data Computing, The Chinese University of Hong Kong, Shenzhen, the Longgang District Special Funds for Science and Technology Innovation (LGKCSDPT2023002), the Guangxi Key R&D Project (No. AB24010167), the Hospital University United Fund of The Second Affiliated Hospital, School of Medicine, The Chinese University of Hong Kong, Shenzhen (No. HUUF-MS-202303), the National Natural Science Foundation of China (NO. 62322608), Guangdong Basic and Applied Basic Research Foundation under Grant 2024A1515010255.

Disclosure of Interests. All authors have no competing interests to declare.

References

1. Ali, S., Ghatwary, N., Jha, D., Isik-Polat, E., Polat, G., Yang, C., Li, W., Galdran, A., Ballester, M.A.G., Thambawita, V., et al.: Assessing generalisability of deep learning-based polyp detection and segmentation methods through a computer vision challenge. arXiv preprint arXiv:2202.12031 (2022)
2. Ali, S., Ghatwary, N., Jha, D., Isik-Polat, E., Polat, G., Yang, C., Li, W., Galdran, A., Ballester, M.Á.G., Thambawita, V., et al.: Assessing generalisability of deep learning-based polyp detection and segmentation methods through a computer vision challenge. arXiv preprint arXiv:2202.12031 (2022)
3. Duan, K., Bai, S., Xie, L., Qi, H., Huang, Q., Tian, Q.: Centernet: Keypoint triplets for object detection. In: Proceedings of the IEEE/CVF international conference on computer vision. pp. 6569–6578 (2019)
4. null Endoscopic Classification Review Group, et al.: Update on the paris classification of superficial neoplastic lesions in the digestive tract. Endoscopy **37**(06), 570–578 (2005)
5. Inoue, H., Kashida, H., ei Kudo, S., Sasako, M., Shimoda, T., Watanabe, H., Yoshida, S., Guelrud, M.B., Lightdale, C.J., Wang, K.K., Riddell, R.H., Diébold, Lambert, R., Rey, J.F., Jung, M.E., Neuhaus, H., Axon, A.T.R., Genta, R.M., Gonvers, J.J.: The paris endoscopic classification of superficial neoplastic lesions: esophagus, stomach, and colon: November 30 to december 1, 2002. Gastrointestinal endoscopy **58 6 Suppl**, S3–43 (2003), https://api.semanticscholar.org/CorpusID:36658734

6. Jiang, Y., Shi, L., Qi, W., Chen, L., Li, G., Han, X., Wan, X., Liu, S.: Automatic bleeding risk rating system of gastric varices. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 3–12. Springer (2023)
7. Jiang, Y., Zhang, Z., Hu, Y., Li, G., Wan, X., Wu, S., Cui, S., Huang, S., Li, Z.: Ecc-polypdet: Enhanced centernet with contrastive learning for automatic polyp detection. IEEE Journal of Biomedical and Health Informatics (2023)
8. Lin, T.Y., Goyal, P., Girshick, R., He, K., Dollár, P.: Focal loss for dense object detection. In: Proceedings of the IEEE international conference on computer vision. pp. 2980–2988 (2017)
9. Ren, M., Triantafillou, E., Ravi, S., Snell, J., Swersky, K., Tenenbaum, J.B., Larochelle, H., Zemel, R.S.: Meta-learning for semi-supervised few-shot classification. arXiv preprint arXiv:1803.00676 (2018)
10. Ren, S., He, K., Girshick, R., Sun, J.: Faster r-cnn: Towards real-time object detection with region proposal networks. Advances in neural information processing systems **28** (2015)
11. Tian, Z., Shen, C., Chen, H., He, T.: Fcos: Fully convolutional one-stage object detection. In: Proceedings of the IEEE/CVF international conference on computer vision. pp. 9627–9636 (2019)
12. Wang, X., Huang, T., Darrell, T., Gonzalez, J., Yu, F.: Frustratingly simple few-shot object detection. arxiv 2020. arXiv preprint arXiv:2003.06957
13. Wu, L., Hu, Z., Ji, Y., Luo, P., Zhang, S.: Multi-frame collaboration for effective endoscopic video polyp detection via spatial-temporal feature transformation. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part V 24. pp. 302–312. Springer (2021)
14. Yan, X., Chen, Z., Xu, A., Wang, X., Liang, X., Lin, L.: Meta r-cnn: Towards general solver for instance-level low-shot learning. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 9577–9586 (2019)
15. Zhang, H., Li, F., Liu, S., Zhang, L., Su, H., Zhu, J., Ni, L.M., Shum, H.Y.: Dino: Detr with improved denoising anchor boxes for end-to-end object detection. arXiv preprint arXiv:2203.03605 (2022)
16. Zhang, X., Yin, K., Liu, S., Feng, Z., Han, X., Li, G., Wan, X.: Self-and semi-supervised learning for gastroscopic lesion detection. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 83–93. Springer (2023)