GeoGCD: Geometry-Aware Prototype Learning for Histopathology Generalized Cancer Category Discovery

Anonymous authors
Paper under double-blind review

Abstract

Generalized Category Discovery (GCD) aims to simultaneously recognize labeled categories and discover novel categories from partially annotated data. Histopathology generalized category discovery is particularly challenging due to the subtle morphological differences among cancer subtypes and the substantial appearance variations within the same subtype, which often lead to ambiguous decision boundaries and less discriminative feature representations. Existing GCD methods mainly focus on classifier optimization or clustering objectives while paying limited attention to preserving fine-grained histopathological structures and exploiting the geometric relationships among category prototypes.To address these challenges, we propose GeoGCD, a geometry-aware prototype learning framework for histopathology generalized category discovery. Specifically, we introduce a Feature Detail Refinement and Aggregation (FDRA) module that enhances discriminative local structural representations by preserving fine-grained pathological details while suppressing redundant feature responses. Furthermore, we develop an Adaptive Margin Prototype Adapter (AMPA) module that dynamically calibrates classification margins according to the geometric relationships among category prototypes, encouraging better discrimination between confusing cancer subtypes. By jointly improving local structural representations and prototype distributions, GeoGCD learns more compact intra-class representations and better-separated inter-class prototypes, leading to more reliable novel category discovery.Extensive experiments on multiple histopathology benchmark datasets demonstrate that GeoGCD consistently outperforms existing generalized category discovery methods under realistic semi-supervised settings, validating its effectiveness for discovering previously unseen cancer categories.

1 Introduction

Cancer subtype discovery from histopathology images is crucial for precision diagnosis and personalized treatment planning Litjens et al. (2017); Komura & Ishikawa (2018). However, large-scale annotation requires substantial clinical expertise and labor Campanella et al. (2019); Bulten et al. (2020), while unlabeled clinical data may contain previously unseen or rare cancer subtypes. Generalized Category Discovery (GCD) Vaze et al. (2022) addresses this setting by simultaneously recognizing labeled categories and discovering novel categories from partially annotated data, providing a scalable paradigm for histopathology analysis.

Existing GCD methods can be broadly categorized into parametric Wen et al. (2023); Guo et al. (2025) and non-parametric Vaze et al. (2022); Pu et al. (2023) approaches. Parametric methods jointly optimize classifiers and feature representations but may exhibit bias toward known categories, whereas non-parametric methods formulate category discovery as clustering and are sensitive to pseudo-label quality and overlapping feature distributions. Moreover, most existing GCD methods are developed for natural images and do not explicitly account for the distinctive characteristics of histopathology images. As shown in Fig. 1, different cancer subtypes may exhibit highly similar tissue architectures, while samples within the

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

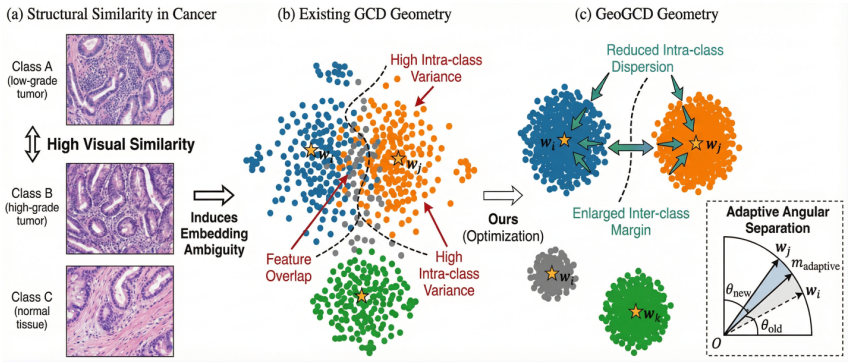

Figure 1: Challenges of histopathological GCD and our GeoGCD framework. (a) Fine-grained similarity between cancer subtypes induces embedding ambiguity. (b) Existing GCD methods generate overlapping features with loose intra-class and insufficient inter-class separation. (c) GeoGCD leverages FDRA and AMPA to optimize local structure and prototype distributions for compact intra-class features and discriminative prototypes.

same subtype can show substantial appearance variations caused by staining, tissue heterogeneity, and imaging conditions Komura & Ishikawa (2018); Lu et al. (2021). These factors lead to entangled feature distributions, large intra-class dispersion, and ambiguous category boundaries, challenging both known-category recognition and novel-category discovery.

To tackle these challenges, we propose GeoGCD, a geometry-aware prototype learning framework for histopathological GCD. GeoGCD optimizes the embedding space via two complementary designs. First, the Feature Detail Refinement and Aggregation (FDRA) module preserves fine-grained tissue structures through frequency-aware enhancement and feature aggregation to yield compact, discriminative representations. Second, the Adaptive Margin Prototype Adapter (AMPA) dynamically calibrates prototype classification margins based on geometric relationships, effectively separating confusing cancer subtypes. Jointly refining feature representations and prototype geometry enables GeoGCD to improve known-category recognition and novel-category discovery.

Under semi-supervised optimization, contrastive learning Chen et al. (2020); Khosla et al. (2020) and pseudo-label guidance Vaze et al. (2022); Wen et al. (2023) learn refined features and geometry-aware prototypes within the GCD framework. Experiments on histopathology benchmarks validate GeoGCD for seen- and novel-category tasks.

The main contributions of this work are summarized as follows:

- We identify the distinctive challenges of histopathology GCD arising from fine-grained structural similarity and substantial intra-class appearance variation, highlighting the importance of discriminative representations and prototype geometry.
- We propose **GeoGCD**, a geometry-aware prototype learning framework that jointly improves feature representations and prototype distributions for histopathology GCD.
- We introduce **FDRA** for fine-grained structural representation enhancement and **AMPA** for geometry-aware adaptive margin optimization, improving intra-class compactness and inter-class separability.

2 Related work

2.1 Parametric methods:

Parametric GCD methods jointly optimize classifiers and feature representations to recognize known categories and discover novel ones. Representative approaches, including GCD Vaze et al. (2022), SimGCD Wen et al. (2023), UNO Fini et al. (2021), and Prompt-

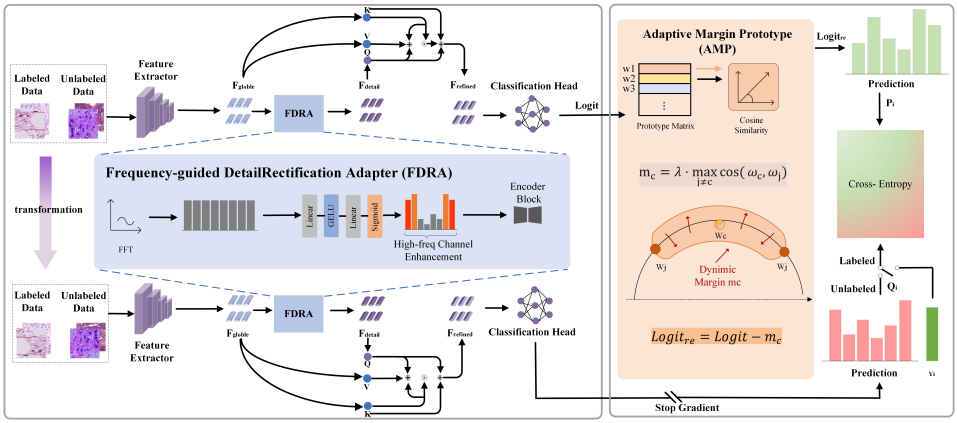

Figure 2: Overview of GeoGCD. Labeled and unlabeled data are processed by the FDRA module for feature refinement and a dual-head structure for classification and representation learning. AMPA adaptively adjusts classification margins according to prototype geometry.

CAL Wang et al. (2024), combine supervised learning with contrastive learning or pseudo-label optimization. Recent methods, such as ProtoGCD Ma et al. (2025), Debiased GCD Guo et al. (2025), MSGCD Duan et al. (2025), and MedGCD Das et al. (2025), further incorporate prototype learning, debiasing, or medical-specific constraints. Adversarial Zhou & Patel (2022) and neighborhood-aware learning Yang et al. (2025); Yu et al. (2023); Du et al. (2023) have also been explored to improve representation robustness. However, these methods primarily focus on classifier optimization or sample relationships and pay limited attention to fine-grained structural representations, which are critical for distinguishing visually similar histopathology subtypes.

2.2 Non-parametric methods:

Non-parametric GCD methods formulate category discovery as clustering without explicitly learning parametric classifiers. Representative approaches include semi-supervised k-means AlZuhair et al. (2025), spectral clustering Bai et al. (2023), graph-based clustering Gao et al. (2023), and density-guided clustering Xu et al. (2023). While these methods reduce classifier bias, their performance depends heavily on feature quality, and overlapping histopathological feature distributions can lead to unreliable clustering and pseudo-labels.

In contrast, GeoGCD jointly improves structural representations and prototype geometry. FDRA enhances fine-grained histopathological features through frequency-aware representation learning, while AMPA adaptively calibrates classification margins according to prototype relationships, yielding a more discriminative embedding space for histopathology GCD.

3 Method

3.1 Problem Definition

Given a labeled dataset $\mathcal{D}_l = \{(x_i, y_i)\}_{i=1}^{N_l}$, where $y_i \in C_k$, and an unlabeled dataset $\mathcal{D}_u = \{x_j\}_{j=1}^{N_u}$ containing samples from known classes C_k and novel classes C_n, Generalized Category Discovery (GCD) Vaze et al. (2022) aims to recognize known categories and discover novel ones. Here, $C_k \cap C_n = \varnothing$ and $C = C_k \cup C_n$. Only $\mathcal{D}_l$ is annotated during training, and the learned model is evaluated on both known and novel categories.

3.2 Overview

GeoGCD jointly enhances fine-grained representations and prototype-aware decision boundaries through two components: Feature Detail Refinement and Aggregation (FDRA) and Adaptive Margin Prototype Adapter (AMPA).Given an input image, the backbone produces a global representation $\mathbf{f}^{global}$, which is refined by FDRA through frequency-aware detail enhancement and cross-attention, yielding $\mathbf{f}^{refined}$. AMPA then adjusts the target-class margin according to prototype geometry, while unlabeled samples follow the original GCD self-training paradigm. Thus, GeoGCD modifies the representation and classification stages while preserving the original semi-supervised framework.

3.3 Geometry-aware Representation Learning

Histopathology images contain subtle morphological structures that are important for cancer subtype discrimination Komura & Ishikawa (2018); Lu et al. (2021). However, global feature aggregation may weaken such fine-grained cues Dosovitskiy et al. (2021); Caron et al. (2021); Wen et al. (2023). We therefore enhance local structural details while preserving global semantics through frequency-guided detail rectification and cross-attention aggregation.

3.3.1 Frequency-guided Detail Rectification

Fine-grained histopathological patterns, including gland boundaries, nuclear arrangements, and tissue textures, are often associated with high-frequency responses and may be weakened during feature extraction Qin et al. (2021b;a). We therefore exploit frequency-domain information to enhance morphology-sensitive representations.

Given the global feature map $\mathbf{f}^{global} \in \mathbb{R}^{C\times H\times W}$, we first apply the two-dimensional Fast Fourier Transform (FFT):

$$\mathcal{F} = \mathrm{FFT}(\mathbf{f}^{global}). \tag{1}$$

The frequency representation is decomposed into low- and high-frequency components using complementary masks:

$$\mathcal{F}_{low} = \mathcal{F} \odot \mathbf{M}_{low}, \qquad \mathcal{F}_{high} = \mathcal{F} \odot \mathbf{M}_{high}, \tag{2}$$

where $\odot$ denotes element-wise multiplication and

$$\mathbf{M}_{high} = \mathbf{1} - \mathbf{M}_{low}. \tag{3}$$

The low-frequency component preserves slowly varying structural information, while the high-frequency component captures local details. We transform the latter back to the spatial domain:

$$\mathbf{f}^{high} = \mathrm{IFFT}(\mathcal{F}_{high}), \tag{4}$$

where $\mathrm{IFFT}(\cdot)$ denotes the inverse Fourier transform.

We then use a lightweight frequency-aware gate to select informative structural responses:

$$\mathbf{g} = \sigma\left(\mathrm{Linear}_2\left(\mathrm{GELU}\left(\mathrm{Linear}_1\left(\mathrm{Pool}(\mathbf{f}^{high})\right)\right)\right)\right), \tag{5}$$

where $\sigma(\cdot)$ denotes the sigmoid activation and $\mathbf{g}$ is the channel-wise structural attention weight. The global feature is then recalibrated as

$$\mathbf{f}^{detail} = \mathrm{Encoder}\left(\mathbf{g} \odot \mathbf{f}^{global}\right), \tag{6}$$

yielding a detail-enhanced representation.

3.3.2 Cross-attention Feature Aggregation

The detail representation may lack sufficient global context. We therefore use cross-attention to integrate global and structural information. The global feature serves as the query, while the detail feature serves as the key and value:

$$\mathrm{Attention}(\mathbf{Q},\mathbf{K},\mathbf{V}) = \mathrm{Softmax}\left(\frac{\mathbf{Q}\mathbf{K}^{\top}}{\sqrt{d_k}}\right)\mathbf{V}, \tag{7}$$

where d_k is the key dimension, with

$$\mathbf{Q} = \mathbf{f}^{global}, \qquad \mathbf{K} = \mathbf{V} = \mathbf{f}^{detail}. \tag{8}$$

The refined representation is

$$\mathbf{f}^{refined} = \text{Attention}\left(\mathbf{f}^{global}, \mathbf{f}^{detail}, \mathbf{f}^{detail}\right). \tag{9}$$

This aggregation preserves global semantics while incorporating fine-grained morphological information for prototype-based classification and category discovery.

3.4 Geometry-aware Prototype Learning

Existing GCD methods often employ shared or fixed classification margins Deng et al. (2019); Wang et al. (2018); Wen et al. (2023); Ma et al. (2025), overlooking different inter-class similarities. We therefore derive adaptive margins from prototype geometry to construct class-dependent decision boundaries.

3.4.1 Prototype Relationship Modeling

Let $\mathbf{w}_c \in \mathbb{R}^D$ denote the classifier prototype of category c. We normalize each prototype as

$$\hat{\mathbf{w}}_c = \frac{\mathbf{w}_c}{\|\mathbf{w}_c\|_2}. \tag{10}$$

The normalized prototype matrix is

$$\mathbf{W} = [\hat{\mathbf{w}}_1, \hat{\mathbf{w}}_2, \cdots, \hat{\mathbf{w}}_C], \tag{11}$$

where $C = |C|$ denotes the total number of categories. The prototype directly corresponds to the class-wise classifier direction.

We measure the relationship between two categories using cosine similarity:

$$\gamma_{c,k} = \hat{\mathbf{w}}_c^\top \hat{\mathbf{w}}_k. \tag{12}$$

A larger $\gamma_{c,k}$ indicates greater prototype similarity and potential category ambiguity. For each category, we define

$$a_c = \max_{k \neq c} \gamma_{c,k}. \tag{13}$$

Thus, a larger a_c indicates a closer neighboring prototype and a greater need for separation.

3.4.2 Adaptive Margin Optimization

Based on prototype ambiguity, AMPA constructs a class-dependent margin:

$$m_c = \lambda_m a_c, \tag{14}$$

where λ_m controls the margin magnitude and is distinct from the loss-balancing coefficient λ.

Let $\mathbf{z}_i$ denote the normalized feature of sample i. The classification logit for category c is

$$s_{ic} = \frac{\mathbf{z}_i^\top \hat{\mathbf{w}}_c}{\tau_c}, \tag{15}$$

where τ_c is the classification temperature.

For a labeled sample with ground-truth category y_i, AMPA modifies only the target-class logit:

$$\tilde{s}_{ic} = s_{ic} - \mathbb{I}(c = y_i) m_{y_i}, \tag{16}$$

where $\mathbb{I}(\cdot)$ is the indicator function. Equivalently,

$$\tilde{s}_{iy_i} = s_{iy_i} - m_{y_i}, \qquad \tilde{s}_{ic} = s_{ic}, \quad c \neq y_i. \tag{17}$$

The supervised classification loss is

$$\mathcal{L}_{cls}^s = -\frac{1}{|B_l|} \sum_{i \in B_l} \log \frac{\exp(\tilde{s}_{iy_i})}{\sum_{c=1}^{C} \exp(\tilde{s}_{ic})}, \tag{18}$$

where B_l denotes the labeled mini-batch. Hence, categories with larger prototype ambiguity require stronger target-class evidence for correct classification, yielding geometry-aware decision boundaries.

3.5 Semi-supervised Optimization

Following SimGCD Wen et al. (2023), GeoGCD adopts its semi-supervised optimization pipeline with no extra standalone loss. FDRA refines features prior to representation learning, and AMPA adjusts logits before classification. The total objective reads:

$$\mathcal{L} = \mathcal{L}_{rep} + \mathcal{L}_{cls}. \tag{19}$$

The representation loss $\mathcal{L}_{rep}$ combines unsupervised and supervised contrastive terms:

$$\mathcal{L}^u_{rep} = -\frac{1}{|B|}\sum_{i\in B}\log\frac{\exp(\mathbf{z}_i^\top \mathbf{z}'_i/\tau_u)}{\sum_{n\in B}\exp(\mathbf{z}_i^\top \mathbf{z}'_n/\tau_u)}, \tag{20}$$

where $\mathbf{z}_i, \mathbf{z}'_i$ denote normalized embeddings and τ_u is the unsupervised temperature. For labeled data, supervised contrastive loss treats same-class samples as positives:

$$\mathcal{L}^s_{rep} = -\frac{1}{|B_l|}\sum_{i\in B_l}\frac{1}{|N_i|}\sum_{q\in N_i}\log\frac{\exp(\mathbf{z}_i^\top \mathbf{z}'_q/\tau_s)}{\sum_{n\in B_l}\exp(\mathbf{z}_i^\top \mathbf{z}'_n/\tau_s)}, \tag{21}$$

where N_i contains samples sharing the class of x_i and τ_s is the supervised temperature. We combine them as

$$\mathcal{L}_{rep} = (1-\lambda)\mathcal{L}^u_{rep} + \lambda\mathcal{L}^s_{rep}, \tag{22}$$

with λ balancing the two terms.

For prototype classification, we follow SimGCD and merge supervised loss and pseudo-label loss for unlabeled samples:

$$\begin{aligned}\mathcal{L}^u_{cls} &= \frac{1}{|B|}\sum_{i\in B}\ell(q'_i, p_i) - \varepsilon H(\bar{p}),\\ \mathcal{L}^s_{cls} &= \frac{1}{|B_l|}\sum_{i\in B_l}\ell(y_i, p_i),\end{aligned} \tag{23}$$

where q'_i is the detached target, p_i the model prediction, $\bar{p}$ batch-averaged prediction, and ε weights entropy regularization. For labeled samples, p_i is computed from AMPA-adjusted logits $\tilde{s}_{ic}$ in Eq. equation 16. The classification objective is

$$\mathcal{L}_{cls} = (1-\lambda)\mathcal{L}^u_{cls} + \lambda\mathcal{L}^s_{cls}. \tag{24}$$

FDRA and AMPA are jointly trained together with the original GCD losses. FDRA strengthens fine-grained features, and AMPA calibrates logits based on prototype ambiguity. This design retains the semi-supervised pipeline while improving feature and classification geometry.

4 Experiments

4.1 Experimental Setup

We evaluate GeoGCD on three public histopathology datasets: LC25000 Borkowski et al. (2019), BreakHis Spanhol et al. (2016), and BACH Aresta et al. (2019). Following the standard GCD protocol Vaze et al. (2022); Wen et al. (2023), each dataset is split into seen and novel categories under semi-supervised settings. We report clustering accuracy (ACC) with Hungarian matching, including overall (All), seen-category (Seen), and novel-category (Novel) accuracy. We adopt a DINO-pretrained ViT-B/16 Caron et al. (2021), using the `[CLS]` token as image representation and fine-tuning only the last transformer block as in SimGCD Wen et al. (2023). Models are trained for 200 epochs with batch size 128 via SGD under cosine learning rate scheduling, with initial lr 0.001. We set $\lambda = 0.35$, $\tau_u = 0.07$, $\tau_c = 1.0$, $\tau_s = 0.1$; the teacher temperature linearly decreases from 0.07 to 0.04 in the first 30 epochs. All baselines share the same backbone and training pipeline for fair comparison.

Table 1: Results on the pathological image datasets. All, Seen, and Novel denote the overall, seen-category, and novel-category accuracy (%), respectively. The best results are shown in bold.

Method	LC			BreakHis			BACH		
	All	Seen	Novel	All	Seen	Novel	All	Seen	Novel
GCD Vaze et al. (2022)	45.8	55.9	45.7	54.3	66.8	43.3	40.3	35.9	47.9
SimGCD Wen et al. (2023)	72.3	75.9	96.8	66.4	81.1	67.1	73.8	76.4	69.2
PCR Duan et al. (2024)	67.8	65.7	70.1	59.8	67.1	45.3	45.6	39.1	50.4
DCR Hu et al. (2025)	71.2	73.7	97.6	52.9	53.1	32.7	56.4	57.1	60.0
ProtoGCD Ma et al. (2025)	70.1	70.2	96.5	61.5	80.4	45.6	70.2	70.4	69.2
MedGCD Das et al. (2025)	70.5	69.5	97.1	62.4	81.4	65.4	72.1	73.6	70.2
Ours	**72.6**	**76.1**	**97.7**	**67.2**	**81.9**	**69.0**	**74.1**	**76.9**	**70.4**

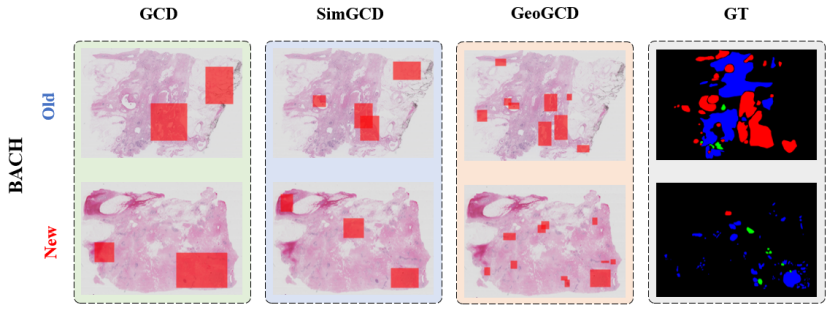

Figure 3: Grad-CAM visualization on the BACH dataset.

4.2 Comparison with State-of-the-art

4.2.1 Quantitative Comparison

As shown in Table 1, GeoGCD achieves SOTA performance on three histopathology datasets. It reaches 72.6% on LC, surpassing SimGCD, ProtoGCD and MedGCD. On BreakHis, our method obtains 67.2% overall and 69.0% novel accuracy (0.8% and 1.9% higher than SimGCD). On BACH, it attains 74.1% overall and 70.4% novel accuracy. This confirms GeoGCD strengthens representation discriminability for both seen- and novel-class tasks.

4.2.2 Qualitative Comparison

We perform qualitative evaluation via attention maps and feature embedding visualization. Grad-CAM is used to examine model localization, and t-SNE to analyze representation discriminability.

Grad-CAM Visualization. Fig. 3 compares Grad-CAM results of GeoGCD and representative GCD methods. GeoGCD accurately focuses on diagnostic regions and suppresses background noise. Compared with GCD and SimGCD, its activation maps highlight pathological areas more clearly and retain fine-grained structures, matching FDRA's design for structural representation enhancement.

Feature Embedding Visualization. We adopt t-SNE to further evaluate representation discriminability. As shown in Fig. 4, GeoGCD yields more compact intra-class clusters and sharper inter-class boundaries than SimGCD. Visually similar cancer subtypes are better separated, and samples within one category cluster tightly. This observation aligns with the discriminative features learned by FDRA and AMPA, constructing a well-structured embedding space for GCD.

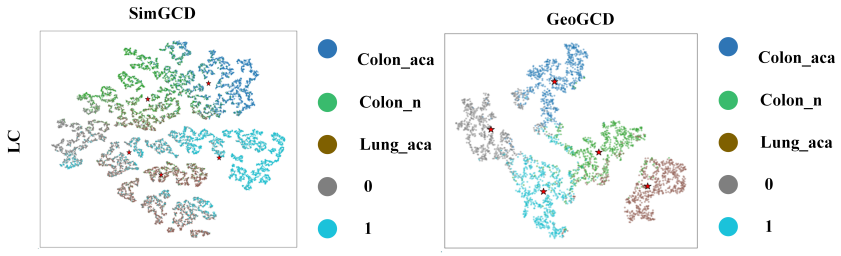

Figure 4: t-SNE visualization of the learned feature embeddings on the LC dataset.

Table 2: Ablation study of FDRA and AMPA. All, Seen, and Novel denote overall, seen-category, and novel-category accuracy (%), respectively.

Components		LC			BreakHis			BACH		
FDRA	AMPA	All	Seen	Novel	All	Seen	Novel	All	Seen	Novel
×	×	72.3	75.9	96.8	66.4	81.1	67.5	73.8	76.4	69.2
✓	×	71.6	76.1	97.1	66.9	81.2	68.1	72.1	75.9	70.1
×	✓	72.1	76.0	97.3	66.8	81.3	68.6	73.9	76.4	70.2
✓	✓	**72.6**	**76.1**	**97.7**	**67.2**	**81.9**	**69.0**	**74.1**	**76.9**	**70.4**

4.3 Ablation Study

4.3.1 Effectiveness of FDRA

To evaluate the effectiveness of FDRA, we compare the baseline with the variant equipped only with FDRA. As shown in Table 2, FDRA mainly improves novel-category performance across the three datasets. Specifically, the novel-category accuracy increases from 96.8% to 97.1% on LC, from 67.5% to 68.1% on BreakHis, and from 69.2% to 70.1% on BACH. These results indicate that FDRA can enhance fine-grained structural representations and facilitate the discovery of novel categories.

4.3.2 Effectiveness of AMPA

We further evaluate the contribution of AMPA by incorporating it alone into the baseline. As shown in Table 2, AMPA consistently improves novel-category accuracy across all three datasets, with gains of 0.5, 1.1, and 1.0 percentage points on LC, BreakHis, and BACH, respectively. The overall accuracy also improves on BreakHis and BACH, while remaining comparable on LC. These results suggest that the geometry-aware adaptive margin particularly benefits the separation of novel categories.

Finally, integrating FDRA and AMPA yields the best performance on all three datasets. GeoGCD achieves overall accuracies of 72.6%, 67.2%, and 74.1% on LC, BreakHis, and BACH, respectively. The full model outperforms the baseline by 0.3, 0.8, and 0.3 percentage points. This verifies that FDRA and AMPA complement each other for representation refinement and prototype discrimination.

4.4 Further Analysis

4.4.1 Feature Geometry Analysis

To further investigate the geometric structure of the learned feature space, we analyze the intra-class and inter-class feature distributions. As shown in Fig. 6(a), GeoGCD achieves the smallest intra-class distance ($D_{\text{intra}} = 0.52$), compared with SimGCD (0.71) and ProtoGCD (0.65), indicating a more compact intra-class distribution. Meanwhile, Fig. 6(b) shows that GeoGCD obtains the largest inter-class distance ($D_{\text{inter}} = 1.56$), compared with SimGCD (1.24) and ProtoGCD (1.31). These results indicate that the combination of FDRA and AMPA produces a more compact intra-class distribution and a more separated inter-class structure, supporting the geometry-aware design of GeoGCD.

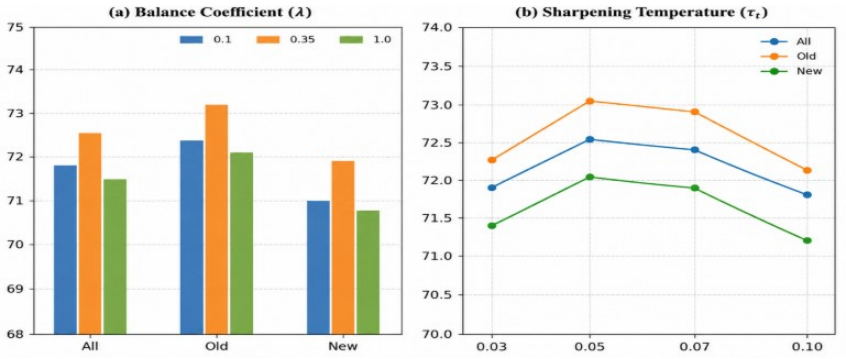

Figure 5: Sensitivity analysis of loss balance coefficient λ and sharpening temperature τ_t.

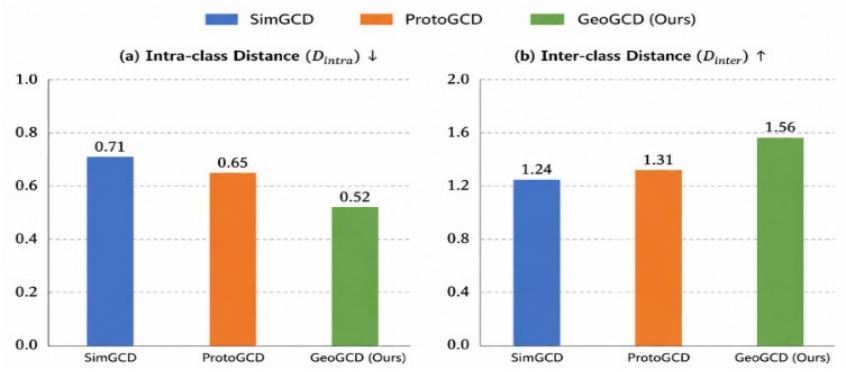

Figure 6: Feature geometry analysis. Comparison of (a) intra-class distance (D_{intra}) and (b) inter-class distance (D_{inter}) for different GCD methods.

4.4.2 Hyperparameter Sensitivity

We analyze the sensitivity of two key hyperparameters: loss balance coefficient λ and sharpening temperature τ_t.

Loss Balance Coefficient λ. As shown in Fig. 5(a), performance rises first and then slightly drops with increasing λ. GeoGCD obtains optimal results at $\lambda = 0.35$, showing that proper balancing of supervised and unsupervised objectives benefits GCD.

Sharpening Temperature τ_t. Fig. 5(b) presents the effect of τ_t. The peak performance occurs at $\tau_t = 0.05$, and overall accuracy stays stable across tested values. Though novel-category accuracy is more sensitive, fluctuations remain mild, demonstrating GeoGCD's robustness to τ_t.

5 Conclusion

We propose GeoGCD, a geometry-aware prototype learning framework for generalized category discovery in histopathology images. By integrating the Frequency-guided Detail Rectification Adapter (FDRA) and the Adaptive Margin Prototype Adapter (AMPA), GeoGCD jointly enhances discriminative representations and prototype separability. Extensive experiments on multiple histopathology datasets demonstrate that GeoGCD consistently outperforms existing GCD methods in both known-category recognition and novel-category discovery.

References

Mona Suliman AlZuhair, Mohamed Maher Ben Ismail, and Ouiem Bchir. Novel dual-constraint-based semi-supervised deep clustering approach. *Sensors*, 25(8):2622, 2025.

Guilherme Aresta, Teresa Araújo, Scotty Kwok, Sai Saketh Chennamsetty, Mohammed Safwan, Varghese Alex, Bahram Marami, Marcel Prastawa, Monica Chan, Michael Donovan, et al. Bach: Grand challenge on breast cancer histology images. *Medical image analysis*, 56:122–139, 2019.

Liang Bai, Minxue Qi, and Jiye Liang. Spectral clustering with robust self-learning constraints. *Artificial Intelligence*, 320:103924, 2023.

Andrew A. Borkowski, Linda A. Bui, Lauren B. Thomas, Catherine P. Wilson, Lauren A. DeLand, and Stephen M. Mastorides. Lung and colon cancer histopathological image dataset (lc25000). *arXiv preprint arXiv:1912.12142*, 2019.

Wouter Bulten et al. Automated deep-learning system for gleason grading of prostate cancer. *The Lancet Oncology*, 21(2):233–241, 2020.

Gabriele Campanella et al. Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. *Nature Medicine*, 25(8):1301–1309, 2019.

Mathilde Caron, Hugo Touvron, Ishan Misra, Hervé Jégou, Julien Mairal, Piotr Bojanowski, and Armand Joulin. Emerging properties in self-supervised vision transformers. In *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 9650–9660, 2021.

Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton. A simple framework for contrastive learning of visual representations. In *ICML*, 2020.

Ankit Das, Chandan Gautam, Pritee Agrawal, Feng Yang, Yong Liu, and Ramasamy Savitha. Medgcd: Generalized category discovery in medical imaging. In *Medical Image Computing and Computer Assisted Intervention – MICCAI 2025*, volume 15965 of *Lecture Notes in Computer Science*, pp. 427–437. Springer Nature Switzerland, 2025. doi: 10.1007/978-3-032-04978-0_41.

Jiankang Deng, Jia Guo, Niannan Xue, and Stefanos Zafeiriou. Arcface: Additive angular margin loss for deep face recognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4690–4699, 2019.

Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, and et al. An image is worth 16x16 words: Transformers for image recognition at scale. In *ICLR*, 2021.

Guowang Du, Lihua Zhou, Zhongxue Li, Lizhen Wang, and Kevin Lü. Neighbor-aware deep multi-view clustering via graph convolutional network. *Information Fusion*, 93:330–343, 2023.

Yu Duan, Junzhi He, Runxin Zhang, Rong Wang, Xuelong Li, and Feiping Nie. Prediction consistency regularization for generalized category discovery. *Information Fusion*, 112: 102547, 2024.

Yu Duan, Zhanxuan Hu, Rong Wang, Zhensheng Sun, Feiping Nie, and Xuelong Li. Mutual-support generalized category discovery. *Information Fusion*, 119:103020, 2025.

Enrico Fini, Enver Sangineto, Stéphane Lathuilière, Zhun Zhong, Moin Nabi, and Elisa Ricci. A unified objective for novel class discovery. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 9284–9292, 2021.

Yuelin Gao, Huirong Li, Yani Zhou, and Yijun Chen. Semi-supervised graph regularized concept factorization with the class-driven constraint for image representation. *AIMS Mathematics*, 8(12):28690–28709, 2023. doi: 10.3934/math.20231468. URL https://doi.org/10.3934/math.20231468.

Pengcheng Guo, Yonghong Song, and Boyu Wang. Towards debiased generalized category discovery. In *Proceedings of the Thirty-Fourth International Joint Conference on Artificial Intelligence (IJCAI)*, pp. 5271–5279, 2025.

Zhanxuan Hu, Yu Duan, Yaming Zhang, Rong Wang, and Feiping Nie. Prototypical classifier with distribution consistency regularization for generalized category discovery: A strong baseline. *Neural Networks*, 182:106908, 2025.

Prannay Khosla, Piotr Teterwak, Chen Wang, Aaron Sarna, Yonglong Tian, Phillip Isola, Aaron Maschinot, Ce Liu, and Dilip Krishnan. Supervised contrastive learning. *Advances in neural information processing systems*, 33:18661–18673, 2020.

Daisuke Komura and Shumpei Ishikawa. Machine learning methods for histopathological image analysis. *Computational and Structural Biotechnology Journal*, 16:34–42, 2018.

Geert Litjens, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra Adiyoso Setio, Francesco Ciompi, Mohsen Ghafoorian, Jeroen Awm Van Der Laak, Bram Van Ginneken, and Clara I Sánchez. A survey on deep learning in medical image analysis. *Medical image analysis*, 42:60–88, 2017.

Ming Y. Lu, Richard J. Chen, Drew Wang, et al. Data-efficient and weakly supervised computational pathology on whole-slide images. *Nature Biomedical Engineering*, 5(6): 555–570, 2021.

Shijie Ma, Fei Zhu, Xu-Yao Zhang, and Cheng-Lin Liu. Protogcd: Unified and unbiased prototype learning for generalized category discovery. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 47(7):6022–6038, 2025. doi: 10.1109/TPAMI.2025.3557502.

Nan Pu, Zhun Zhong, Nicu Sebe, and Elisa Ricci. Discovering novel categories via dual clustering and contrastive learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2023.

Xuebin Qin et al. Ff-parser: Fine-grained feature parsing for human parsing. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021a.

Zequn Qin, Pengyi Zhang, Fei Wu, and Xi Li. Fcanet: Frequency channel attention networks. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 783–792, 2021b.

Fábio A. Spanhol, Luiz S. Oliveira, Caroline Petitjean, and Laurent Heutte. A dataset for breast cancer histopathological image classification. *IEEE Transactions on Biomedical Engineering*, 63(7):1455–1462, 2016. doi: 10.1109/TBME.2015.2496264.

Sagar Vaze, Kai Han, Andrea Vedaldi, and Andrew Zisserman. Generalized category discovery. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 7492–7501, 2022.

Hao Wang, Yitong Wang, Zheng Zhou, Xing Ji, Dihong Gong, Jingchao Zhou, Zhifeng Li, and Wei Liu. Cosface: Large margin cosine loss for deep face recognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5265–5274, 2018.

Yikai Wang, Yabin Huang, Kai Zhang, Yifan Wang, and Jiashi Feng. Promptcal: Contrastive affinity learning via auxiliary prompts for generalized category discovery. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024.

Xin Wen, Bingchen Zhao, and Xiaojuan Qi. Parametric classification for generalized category discovery: A baseline study. In *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 16590–16600, 2023.

Xiao Xu, Haiwei Hou, and Shifei Ding. Semi-supervised deep density clustering. *Applied Soft Computing*, 148:110903, 2023. doi: 10.1016/j.asoc.2023.110903. URL https://doi.org/10.1016/j.asoc.2023.110903.

Bin Yang, Leilei Zhong, Xu Cheng, Ying Xing, Zhipu Xie, Jinchao Huang, Yuhao Gao, and Lexi Xu. Neighbor-aware graph representation learning for robust telecom fraud detection. In *2025 IEEE International Conference on High Performance Computing and Communications (HPCC)*, pp. 1–6. IEEE, 2025.

Chunlin Yu, Ye Shi, and Jingya Wang. Contextually affinitive neighborhood refinery for deep clustering. *Advances in Neural Information Processing Systems*, 36:5778–5790, 2023.

Mo Zhou and Vishal M Patel. Enhancing adversarial robustness for deep metric learning. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 15325–15334, 2022.