Anomaly as Non-Conformity via Training-Free Graph Laplacian Energy Minimization

Supplementary Material

Table of Contents

S1. Implementation details

Backbone, preprocessing, and post-processing. We use DINOv3-L/16 [25] as a frozen visual backbone and extract patch-level features from the 18-th Transformer block. All input images are resized to 768×768 and normalized with mean $(0.485, 0.456, 0.406)$ and standard deviation $(0.229, 0.224, 0.225)$. The resulting anomaly map is further smoothed with a Gaussian filter with kernel size 7 and $\sigma = 0.8$.

Inference-time augmentation. At test time, we apply geometric augmentation to both the query image and the few-shot reference images. For each reference image, we retain the original view and generate additional transformed views using random horizontal flipping, random vertical flipping, and an affine transformation consisting of a rotation sampled from $[-15°, -5°] \cup [5°, 15°]$, a translation within $\pm 2\%$ of the image width and height, isotropic scaling in $[0.95, 1.05]$, and shear sampled from $[-5°, 5°]$ along both axes. We generate 25 paired augmented views and 5 additional reference-only augmented views for each reference image. For the paired views, the same transformation parameters are applied to both the query image and the corresponding reference image, whereas the reference-only views are used only to enrich the reference patch pool. During inference, anomaly maps from the augmented query views are mapped back to the original image coordinates and aggregated using entropy-based confidence weighting.

Memory usage. In the 1-shot setting on a single NVIDIA RTX 3090 GPU, the maximum memory usage was approximately 10 GB when augmentation was enabled and 2 GB when augmentation was disabled.

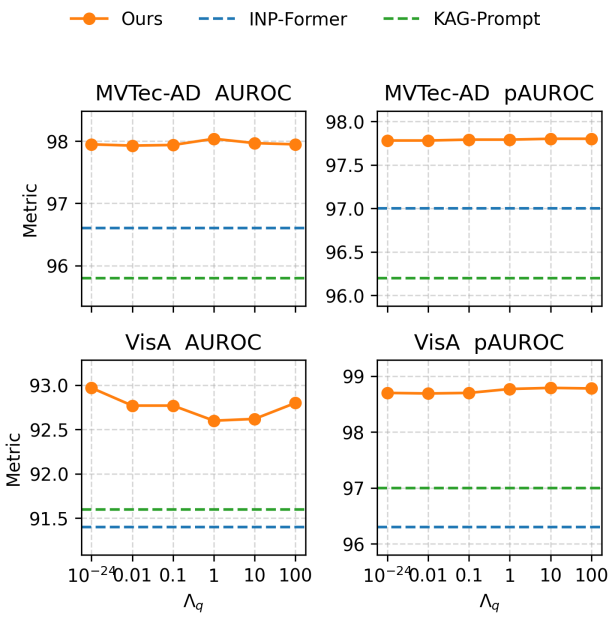

Figure S1. **Performance trend of ANoCo with respect to Λ_q in the 1-shot setting.**

S2. Ablation on the query feature stabilization coefficient Λ_q

We analyze the effect of the query feature stabilization coefficient Λ_q in Figure S1. We sweep $\Lambda_q \in \{10^{-24}, 0.01, 0.1, 1, 10, 100\}$ and observe that ANoCo consistently outperforms all baselines across all values of Λ_q on every dataset. The overall performance curves are relatively flat, indicating that ANoCo is robust to the exact choice of Λ_q and does not require delicate tuning of this coefficient. At the same time, some datasets exhibit small additional gains for moderate values of Λ_q (e.g., around $\Lambda_q = 1$), suggesting that query feature stabilization can be exploited as an optional degree of freedom when further fine-tuning is desired.

S3. Graph update strategies

Table S1 presents an ablation study on graph update strategies under the 1-shot setting. We compare the proposed query optimization via energy minimization with a training-free message-passing alternative. For message passing, we perform iterative propagation using cosine-similarity-based edge weights, which are identical to those used in

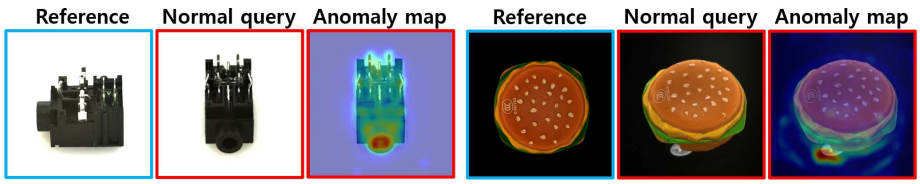

Figure S2. **Failure cases on Real-IAD in the 1-shot setting.**

Table S1. **Ablation study on graph update strategies under the 1-shot setting**. We compare message passing with our energy minimization approach.

Round	MVTec-AD		VisA	
	AUROC	pAUROC	AUROC	pAUROC
1	87.7	93.7	67.5	95.3
2	82.7	92.2	65.1	93.3
3	82.4	92.2	65.6	93.3
4	82.4	92.2	65.3	93.3
5	82.4	92.2	65.6	93.2
ANoCo (Ours)	**97.9**	**97.7**	**92.7**	**98.7**

the energy-minimization formulation.

We evaluate message passing with 1 to 5 propagation rounds. As the number of rounds increases, message passing does not yield consistent improvements in anomaly discrimination, making the choice of the optimal number of rounds non-trivial. In contrast, the energy-minimization strategy of ANoCo achieves consistently stronger and more stable performance across both datasets than the message-passing variants.

Table S2. Ablation study of various ANoCo non-conformity metrics under the 1-shot setting.

Metric	MVTec-AD		VisA	
	AUROC	pAUROC	AUROC	pAUROC
L2	97.6	97.6	92.7	98.7
Cos-dis	97.7	97.6	92.4	98.6
Eq. 10	97.9	97.7	92.7	98.7

S4. Additional details of non-conformity as anomaly evidence

We additionally compare three candidates for the patchwise non-conformity energy E_i used as anomaly evidence: pure Euclidean energy $\|\tilde{\mathbf{f}}_i^q - \mathbf{f}_i^q\|_2^2$, pure cosine dissimilarity $(1 - \cos(\tilde{\mathbf{f}}_i^q, \mathbf{f}_i^q))$, and their product $\|\tilde{\mathbf{f}}_i^q - \mathbf{f}_i^q\|_2^2(1 - \cos(\tilde{\mathbf{f}}_i^q, \mathbf{f}_i^q))$ as adopted in Equation 10 in Table S2. All three options consistently outperform the baselines, and the performance differences between them are relatively small. However, the product form provides a small but consistent improvement over using either term alone. This suggests that ANoCo does not rely on a carefully tuned or dataset-specific non-conformity formulation, while the chosen energy yields a slightly stronger separation between normal and anomalous regions.

Table S3. **Comparison of anomaly detection and localization performance on Real-IAD dataset under different few-shot settings.**

# refs	Method	Real-IAD [S1]	
		AUROC	pAUROC
1-shot	SPADE [5]	51.2	59.5
	PatchCore [24]	59.3	89.6
	WinCLIP [12]	**69.4**	91.9
	PromptAD [18]	52.2	84.9
	INP-Former [19]	67.5	**94.9**
	ANoCo (Ours)	65.4	93.4
2-shot	SPADE [5]	50.9	59.5
	PatchCore [24]	63.3	92.0
	WinCLIP [12]	70.9	93.2
	PromptAD [18]	57.7	86.4
	INP-Former [19]	70.6	96.0
	ANoCo (Ours)	**77.3**	**97.1**
4-shot	SPADE [5]	50.8	59.5
	PatchCore [24]	66.0	92.9
	WinCLIP [12]	73.0	93.8
	PromptAD [18]	59.7	86.9
	INP-Former [19]	76.7	97.3
	ANoCo (Ours)	**83.8**	**98.3**

S5. Evaluation on Real-IAD dataset

In this section, we evaluate ANoCo on the Real-IAD [S1] dataset, and report the results in Table S3, complementing

the MVTec-AD and VisA results in the main paper. Real-IAD consists of 30 object categories, with 36,645 normal images for training and a test set of 63,256 normal images and 51,329 anomalous images. ANoCo shows strong performance in the 2-shot and 4-shot settings, significantly outperforming the baseline. We also analyze failure cases and observe that, in the 1-shot setting, ANoCo can misclassify previously unseen but normal variations as anomalies. This issue is particularly pronounced on Real-IAD due to the large viewpoint variation in query images (Figure S2). However, with more reference images, ANoCo can better capture the normal manifold, resulting in strong performance in the 2-shot and 4-shot settings.

Table S4. **Pixel-level AUROC comparison with AD-DINOv3.**

Method	MVTec-AD	VisA
AD-DINOv3 [S2]	91.6	95.6
ANoCo (Ours)	**97.7**	**98.7**

Table S5. **Preliminary generality analysis with WinCLIP on MVTec-AD under the 1-shot setting.**

Method	MVTec-AD	
	AUROC	AP
WinCLIP	93.1	96.5
WinCLIP + **ANoCo (Ours)**	**93.8**	**97.1**

S6. Additional results

Generality. ANoCo is training-free and can be used with different visual feature extractors, but it is not intended as a fully plug-and-play post-processing framework. Rather, it serves as an alternative to retrieval- or similarity-based few-shot anomaly detection methods. As a preliminary observation, Table S5 shows that the proposed graph optimization module can also be applied to another baseline, WinCLIP.

Comparison with AD-DINOv3. AD-DINOv3 [S2] uses both DINOv3 and CLIP and requires training on both anomalous and normal samples from external datasets. In this work, we focus on a setting where only few-shot normal samples within the same dataset are available. Regardless, we compare against the reported results of AD-DINOv3, where ANoCo is shown to outperform in Table S4.

S7. Per-category performance and detailed visualization

Table S6 reports per-category anomaly detection and localization results on MVTec-AD under the 1-shot setting, providing a more fine-grained view beyond the dataset-level averages reported in the main paper. We also include detailed visualizations to further illustrate category-wise behavior and qualitative characteristics.

Table S6. **Per-category anomaly detection and localization performance of ANoCo on MVTec-AD in the 1-shot setting.**

Category	Image-level			Pixel-level		
	AUROC	AUPR	F1-MAX	AUROC	PRO	F1-MAX
Texture						
carpet	99.8	99.9	100.0	99.6	98.7	76.9
grid	100.0	100.0	100.0	99.5	96.7	60.3
leather	100.0	100.0	100.0	99.3	98.2	46.5
tile	100.0	100.0	100.0	98.1	95.8	75.1
wood	98.6	99.5	97.5	97.0	96.1	70.4
Mean	99.6	99.8	99.5	98.7	97.1	65.8
Object						
bottle	99.9	99.9	99.2	98.7	97.0	78.1
cable	96.0	97.9	93.6	97.0	93.6	72.3
capsule	93.8	98.6	93.9	98.8	97.4	57.5
hazelnut	98.2	99.0	97.1	99.6	95.5	81.2
metal nut	100.0	100.0	100.0	95.5	94.8	70.1
pill	96.5	99.2	97.5	95.0	97.8	47.9
screw	89.7	96.8	88.5	98.4	93.5	53.2
toothbrush	100.0	100.0	100.0	99.3	97.2	66.2
transistor	96.7	96.1	88.6	90.9	80.6	48.6
zipper	99.4	99.8	98.7	99.3	97.6	68.8
Mean	97.0	98.7	95.7	97.2	94.5	64.3
Overall Mean	97.9	99.1	96.9	97.7	95.4	64.9

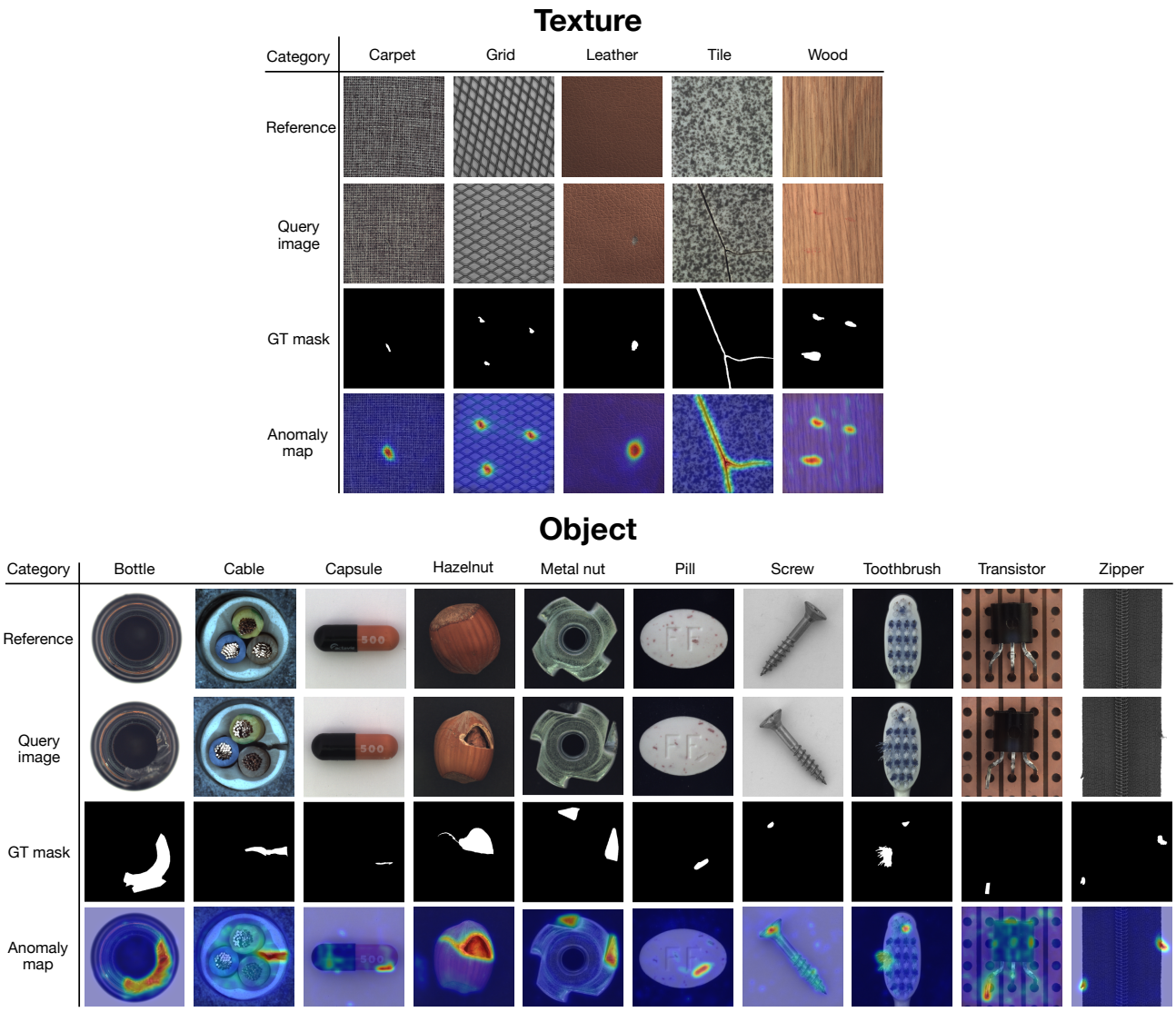

Figure S3. **Qualitative visualization of anomaly localization results on MVTec-AD under the 1-shot setting.**

References

[S1] Chengjie Wang, Wenbing Zhu, Bin-Bin Gao, Zhenye Gan, Jiangning Zhang, Zhihao Gu, Shuguang Qian, Mingang Chen, and Lizhuang Ma. Real-iad: A real-world multi-view dataset for benchmarking versatile industrial anomaly detection. In *CVPR*, 2024. 2

[S2] Jingyi Yuan, Jianxiong Ye, Wenkang Chen, and Chenqiang Gao. Ad-dinov3: Enhancing dinov3 for zero-shot anomaly detection with anomaly-aware calibration. *arXiv preprint arXiv:2509.14084*, 2025. 3