Supplementary Material for ROADS: Robust Prompt-driven Multi-Class Anomaly Detection under Domain Shift

Hossein Kashiani Niloufar Alipour Talemi Fatemeh Afghah
Clemson University
{hkashia, nalipou, fafghah}@clemson.edu

1. Class-wise Results on the VISA Dataset

In this section, we present a detailed analysis of the class-wise robustness performance of our proposed method for both anomaly detection and localization on the VISA dataset [5]. The evaluation is conducted under Gaussian noise distortion as one type of corruption in our paper. We compare the results of five different methods: UniAD [3], ViTAD [4], RD++ [2], SimpleNet [1], and our proposed method, ROADS. Performance metrics are provided for both Image-level AUROC (I-AUROC%) and Pixel-level AUPRO (P-AUPRO%) to offer a comprehensive understanding of each method's efficacy.

1.1. Gaussian Noise Corruption

As shown in Table 1, ROADS achieves significant improvements in both anomaly detection and localization under Gaussian noise corruption across various object classes. In simpler object categories like "Chewing Gum" and "Fryum" ROADS consistently outperforms other methods, demonstrating superior robustness. Specifically, for "Chewing Gum" ROADS reaches an I-AUROC of 96.3%, narrowly outperforming UniAD's 96.2% [3], but far exceeding RD++ [2] and SimpleNet [1], both of which achieve significantly lower scores. Similarly, for "Fryum", ROADS delivers an I-AUROC of 92.2%, outperforming RD++ [2] and SimpleNet [1] by a notable margin.

For more complex categories such as "PCB1", "PCB3" and "PCB4" ROADS maintains high detection accuracy, with I-AUROC scores of 95.8%, 96.0% and 99.7%, respectively. These results further validate ROADS' robustness in out-of-distribution (OOD) settings, as it surpasses both ViTAD [4] and SimpleNet [1] by a significant margin in both cases.

In categories with multiple object instances, such as "Candles" and "Capsules" ROADS excels not only in detection but also achieves substantial improvements in localization accuracy. For example, ROADS achieves an 80.5% P-AUPRO for "Candles" far surpassing SimpleNet's 44.2%, and for "Capsules" it reaches 86.4% P-AUPRO, solidifying its superiority over competing methods.

Overall, ROADS consistently demonstrates strong performance under Gaussian noise, achieving the highest total average I-AUROC (88.6%) and P-AUPRO (84.0%), reinforcing its robustness and effectiveness in handling noisy environments.

2. Heatmaps on the VISA Dataset

To further illustrate the effectiveness of our approach, we present qualitative localization results of anomaly detection for different methods under Gaussian noise corruption in Figure 1. These heatmaps highlight the precision of the detected anomalous regions, showcasing the robustness of ROADS across different methods. Under Gaussian noise corruption, ROADS displays remarkable robustness. For complex object categories such as "Capsules" ROADS produces high-quality anomaly maps with sharper boundaries around the detected regions compared to others. In contrast, the heatmaps from RD++ [2] and UniAD [3] are often less accurate, with instances of mislocalization and reduced confidence in noisy areas. These results underline ROADS' ability to mitigate noise and maintain precise anomaly localization, even under challenging OOD conditions.

References

[1] Zhikang Liu, Yiming Zhou, Yuansheng Xu, and Zilei Wang. Simplenet: A simple network for image anomaly detection and localization. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 20402–20411, 2023. 1, 2

[2] Tran Dinh Tien, Anh Tuan Nguyen, Nguyen Hoang Tran, Ta Duc Huy, Soan Duong, Chanh D Tr Nguyen, and Steven QH Truong. Revisiting reverse distillation for anomaly detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 24511–24520, 2023. 1, 2

[3] Zhiyuan You, Lei Cui, Yujun Shen, Kai Yang, Xin Lu, Yu Zheng, and Xinyi Le. A unified model for multi-class anomaly detection. In *Advances in Neural Information Processing Systems*, pages 4571–4584, 2022. 1, 2

Table 1. Quantitative comparison with SOTA methods on benchmark VISA [5] under Gaussian noise corruption (OOD) settings. Results for anomaly detection and localization are shown as I-AUROC% / P-AUPRO%. The best results are highlighted in bold.

	Category/Method	UniAD [3]	ViTAD [4]	RD++ [2]	SimpleNet [1]	ROADS (Ours)
Single	Cashew	86.2 / 84.6	76.5 / 72.5	72.3 / 84.0	69.6 / 47.7	**92.1** / **85.4**
	Chewing gum	96.2 / **79.4**	87.7 / 68.4	67.9 / 43.4	61.6 / 56.1	**96.3** / 77.3
	Fryum	87.3 / 81.4	**94.8** / 61.9	82.7 / 50.0	65.1 / 73.0	92.2 / **81.7**
	Pipe fryum	**90.5** /81.4	77.6 / 55.1	73.8 / **87.3**	67.5 / 36.4	82.6 / 65.5
Multiple	Candles	82.3 / 79.2	62.5 / 47.5	47.2 / 60.0	54.1 / 44.2	**84.7** / **80.5**
	Capsules	73.2 / 72.7	71.5 / 42.3	73.6 / 77.1	38.8 / 13.7	**83.1** / **86.4**
	Macaroni1	71.6 / 83.7	53.1 / 31.7	57.1 / 79.5	50.4 / 60.9	**73.4** / **87.1**
	Macaroni2	67.4 / 81.2	56.4 / 19.6	48.9 / 83.4	49.6 / 73.6	**71.8** / **92.7**
Complex	PCB1	94.5 / 84.9	80.6 / 46.4	93.6 / 91.8	78.2 / 10.9	**95.8** / **92.5**
	PCB2	91.3 / 80.1	70.9 / 39.9	**97.0** / **88.6**	67.9 / 15.6	95.9 / 84.7
	PCB3	87.7 / 82.0	75.7 / 42.7	88.6 / 88.0	65.7 / 53.2	**96.0** / **88.7**
	PCB4	99.5 / 83.8	98.5 / 77.9	99.1 / 85.1	69.3 / 29.4	**99.7** / **86.5**
Total Average		85.6 / 81.8	75.5 / 50.5	75.1 / 76.5	61.5 / 42.9	**88.6** / **84.0**

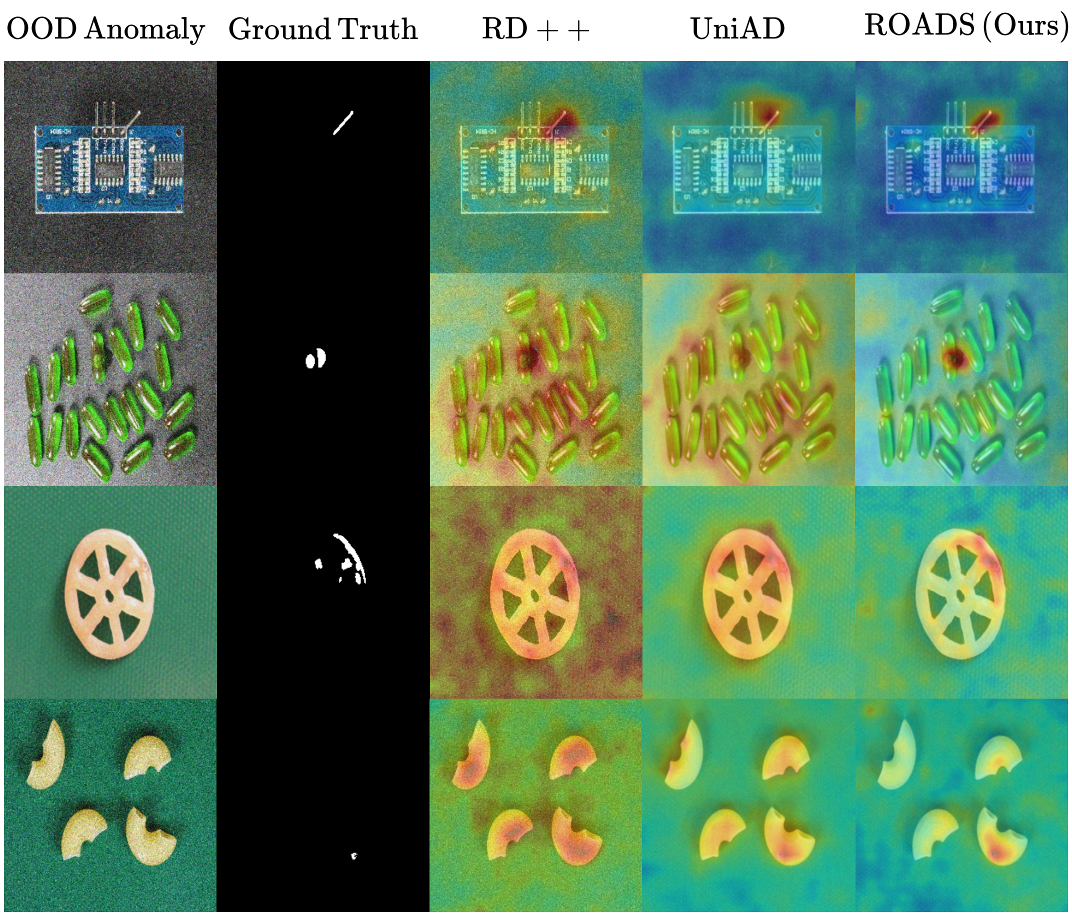

Figure 1. Qualitative comparison between the proposed ROADS method, RD++ [2], and UniAD [3] on the VISA dataset [5] under OOD settings with Gaussian noise.

[4] Jiangning Zhang, Xuhai Chen, Yabiao Wang, Chengjie Wang, Yong Liu, Xiangtai Li, Ming-Hsuan Yang, and Dacheng Tao. Exploring plain ViT reconstruction for multi-class unsupervised anomaly detection. *arXiv preprint arXiv:2312.07495*, 2023. 1, 2

[5] Yang Zou, Jongheon Jeong, Latha Pemula, Dongqing Zhang, and Onkar Dabeer. SPot-the-Difference self-supervised pre-training for anomaly detection and segmentation. In *European Conference on Computer Vision*, pages 392–408. Springer, 2022. 1, 2