Training-Free Industrial Defect Generation with Diffusion Models

Supplementary Material

1. Data Augmentation Strategies

The authenticity of a synthetic defect is determined by both the visual property of the defect and its location. For instance, the defect type "thread top" of screws will appear exclusively on the body of a screw and not appear on the screw tip or head. Enhancing defect diversity while preserving authentic consistency becomes a key challenge in data augmentation during defect generation. To address this, we predefined the object areas where each anomaly type could possibly occur and randomly placed augmented real masks within these regions. The augmentation process involves random scaling, translation, and rotation of the masks, ensuring expanded seen defect distribution without compromising the semantic relevance of their positions.

2. Experiments

2.1. Ablation study

Impact of the number of generated anomalies. We provide an ablation study of generated abnormal ratios as shown in Tab. 1. We investigate the impact of different numbers of synthetic samples on model performances under the general setting. When $N < 500$, the increasing number of generated anomalies can provide the model with more diverse training data, thus enhancing the performance. However, an excessive number of generated samples may lead to over-fitting to specific patterns, reducing the model's ability to generalize on real anomalies. To reach a balance between model performance and the cost of image generation, we select $N = 500$ as the optimal threshold.

Hyperparameter sensitivity. We conduct sensitivity analysis in Tab. 2 by varying guidance strength η, AAM threshold $\tau = \text{mean}(\mathbf{D}_t^{map}) + k \cdot \text{std}(\mathbf{D}_t^{map})$, guidance step T_n, and TP activation step T_i on `carpet` (MVTec AD). The results indicate that our default settings (bolded in the table) provide a favorable balance between stability and fidelity. For instance, $\eta = 0.04$ achieves the best overall performance, while excessively large values of η introduce instability and degrade both Local IS and PRO. Similarly, $k = 0.2$ yields consistent improvements, whereas higher values cause overactivation of anomaly regions, adversely affecting spatial precision and leading to misalignment. For the guidance step T_n, setting it to 30 results in optimal trade-offs. Increasing T_n beyond this point tends to disrupt background recovery, reducing overall image coherence. Regarding T_i, early or delayed activation of the texture preservation module has marginal but observable effects on defect localization. Notably, $T_i = 10$ slightly boosts PRO but compromises overall alignment.

N	i-AUROC	i-AP	p-AUROC	p-AP	Acc
100	99.0	99.6	98.4	78.0	76.8
300	99.2	99.7	98.7	83.2	79.1
500	**99.6**	**99.8**	**99.1**	**84.1**	**79.8**
800	99.2	99.6	98.9	82.6	78.0

Table 1. Impact of the number of generated anomalies.

Parameter setting	Local IS ↑	I-AUC ↑	PRO ↑
$\eta = 0.02$	2.92	96.9	92.8
$\eta = 0.04$ (default)	**3.12**	**98.9**	**96.2**
$\eta = 0.06$	3.03	98.6	94.5
$\eta = 0.08$	2.90	93.0	87.5
$k = 0.0$	3.03	95.3	94.9
$k = 0.2$ (default)	**3.12**	98.9	**96.2**
$k = 0.4$	3.00	**99.1**	95.7
$k = 0.6$	2.99	91.9	90.9
$T_n = 20$	3.03	96.5	95.3
$T_n = 30$ (default)	**3.12**	**98.9**	**96.2**
$T_n = 40$	2.88	97.9	94.9
$T_n = 50$	3.00	94.5	94.3
$T_i = 10$	2.97	95.7	**96.7**
$T_i = 25$ (default)	**3.12**	**98.9**	96.2
$T_i = 40$	2.94	97.2	94.4

Table 2. Ablation study on hyperparameter sensitivity.

2.2. Resource requirement

We perform inference on a single NVIDIA GeForce RTX 3090 24GB GPU, and our experiments demonstrate that our approach is highly resource-efficient. As detailed in Tab. 3, we compare the computational and time costs of mainstream methods. Despite our model's parameter count not being optimal, it requires no additional training for each anomaly type. This unified training approach dramatically reduces overall training time compared to methods like DFMGAN (621 hours) and AnoDiff (390 hours). Moreover, as illustrated in Figure 1, our method exhibits excellent scalability as the number of object categories increases. Both DFMGAN and AnoDiff require retraining or heavy fine-tuning per object type, resulting in a steep linear growth in total generation time. In contrast, our method incurs only a marginal overhead relative to the AnyDoor baseline. Owing to its training-free design and one-shot generalization capability, it exhibits a consistently stable computational cost as the number of object categories increases, demonstrating superior scalability and deployment efficiency in industrial scenarios.

2.3. Extended Application.

In addition to generating defects corresponding to specific object defect types based on real anomalies, our model

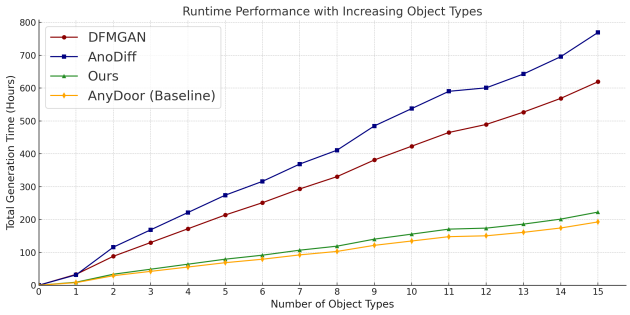


Figure 1. Total generation time on MVTec AD.

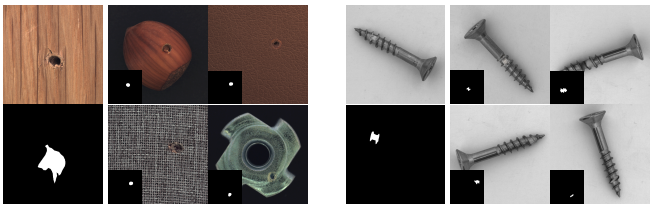
Model	Traning time	Model size	Inference time
DFMGAN	621h	26906M	0.9s
AnoDiff	390h	7240M	20s
Ours	0h	15415M	22s

Table 3. Resource requirement from DFMGAN, AnoDiff and our model on MVTec AD dataset.

demonstrates more comprehensive applicability in practical scenarios. As presented in Fig. 2, the first column on the left side of each section denotes the reference anomalous samples, while the right side displays the outputs generated from various masks.

Cross-object transfer. As shown in Fig. 2(a), our pipeline supports cross-object transfer—transferring defects from one object to another material while achieving seamless background coherence and visual alignment. This capability is particularly promising for tasks in open-set anomaly detection. To further assess the adaptability of our method, we explore its qualitative performance in anomaly transfer, where a hole defect from wood is transferred between different objects while preserving structural integrity and realistic textures of the target object background. As shown in Tab. 4, our method consistently outperforms Crop-Paste and AnoDiff. For hole transfer on carpet, it improves Pro by 10.92% over AnoDiff, while for hazelnut, it achieves a 2.82% gain. These results highlight our method's potential for open-set anomaly detection, effectively synthesizing defects across diverse objects while preserving structural integrity and background coherence.

One anomaly with more mask. We also recognize that the one-shot setting, relying on a single pair of a real anomaly and its corresponding mask, methodologically limits the diversity of the generated anomalies. Building on this perspective, we explore a novel application that enhances model performance by increasing mask diversity while still using only one anomalous sample. As illustrated in Fig. 2(b), additional masks can be conveniently provided through user input. Even with just one real anomaly sample, Our method also significantly enhances the performance of



(a) Quantitative of cross-object transfer (b) Quantitative of one anomaly with more masks

Figure 2. Quantitative examples of extended application.

Defect	Method	i-AUROC	p-AUROC	Pro
Hole transfer on carpet	Crop-Paste	99.40	96.47	90.64
	AnoDiff	96.43	96.65	86.98
	Ours	**100**	**99.72**	**97.90**
Hole transfer on hazenut	Crop-Paste	99.37	94.47	85.75
	AnoDiff	99.16	96.69	94.41
	Ours	**100**	**99.60**	**97.23**

Table 4. Results of anomaly transfer using one anomaly sample.

Setting	i-AUROC	i-AP	p-AUROC	p-AP	Acc
one-shot real masks	88.0	94.0	96.7	32.2	54.3
five-shot real masks	**96.9**	**97.6**	**97.8**	40.3	**85.3**
one anomaly with five masks	90.4	95.3	97.7	**42.7**	82.5

Table 5. Results of generating Screw defects in MVTec AD using one anomaly sample and five masks.

anomaly visual tasks, particularly in the localization and classification sub-tasks (Tab. 5 illustrates an example with screws), with scores closely aligning with the five ground truth masks. This demonstrates that our model can achieve high-quality and diverse anomaly generation, even in conditions with limited anomaly samples, which is highly applicable to real industrial needs.

2.4. More Quantitative Experiments

In this section, we present more comprehensive quantitative results derived from the experiments in our main paper.

Comparision with unsupervised method. In Section 5.3 of the main paper, we integrated our approach with state-of-the-art, lightweight anomaly detection models to conduct comparative experiments. Table 7 illustrates the performance of our method in a one-to-one model framework, where we compare SimpleNet [7], GLAD [8], and GLASS [1]. The results clearly indicate that our method, when combined with GLASS, achieves the highest performance across all metrics, especially on VisA, where it marks an improvement of +0.7% in Pro over GLASS. Furthermore, Table 6 reports the evaluation results for a unified multi-class model, where we compare UniAD [9], DiAD [4], DeSTSeg [10], MambaAD [5], and SimpleNet [7] (trained in a multi-class setting). With the lightweight backbone as UniAD (24.5M parameters, 3.6G FLOPs), our approach improves image-level F1-max by

Method	Params	FLOPs	MVTec AD							VisA						
			I-AUC	I-AP	I-F1-max	P-AUC	P-AP	P-F1-max	Pro	I-AUC	I-AP	I-F1-max	P-AUC	P-AP	P-F1-max	Pro
UniAD [9]	**24.5M**	**3.6G**	96.5	98.8	96.2	96.8	43.4	49.5	90.7	88.8	90.8	85.8	98.3	33.7	39.0	85.5
SimpleNet [7]	72.8M	16.1G	95.3	98.4	95.8	96.9	45.9	49.7	86.5	87.2	87.0	81.8	96.8	34.7	37.8	81.4
DiAD [4]	1525M	451.5G	97.2	99.0	96.5	96.8	52.6	55.5	90.7	86.8	88.3	85.1	96.0	26.1	33.0	75.2
DeSTSeg [10]	35.2M	122.7G	89.2	95.5	91.6	93.1	54.3	50.9	64.8	88.9	89.0	85.2	96.1	39.6	43.4	67.4
MambaAD [5]	25.7M	8.3G	98.6	99.6	97.8	**97.7**	**56.4**	**59.2**	**93.1**	**94.3**	94.5	89.4	98.5	**39.4**	**44.0**	**91.0**
Ours + UniAD	**24.5M**	**3.6G**	**98.9**	**99.7**	**98.1**	97.1	45.2	50.9	91.1	**94.3**	**95.9**	**90.4**	**98.6**	34.1	39.1	86.9

Table 6. Quantitative Results on MVTec AD and VisA for **unified** model between our method and SOTA methods.

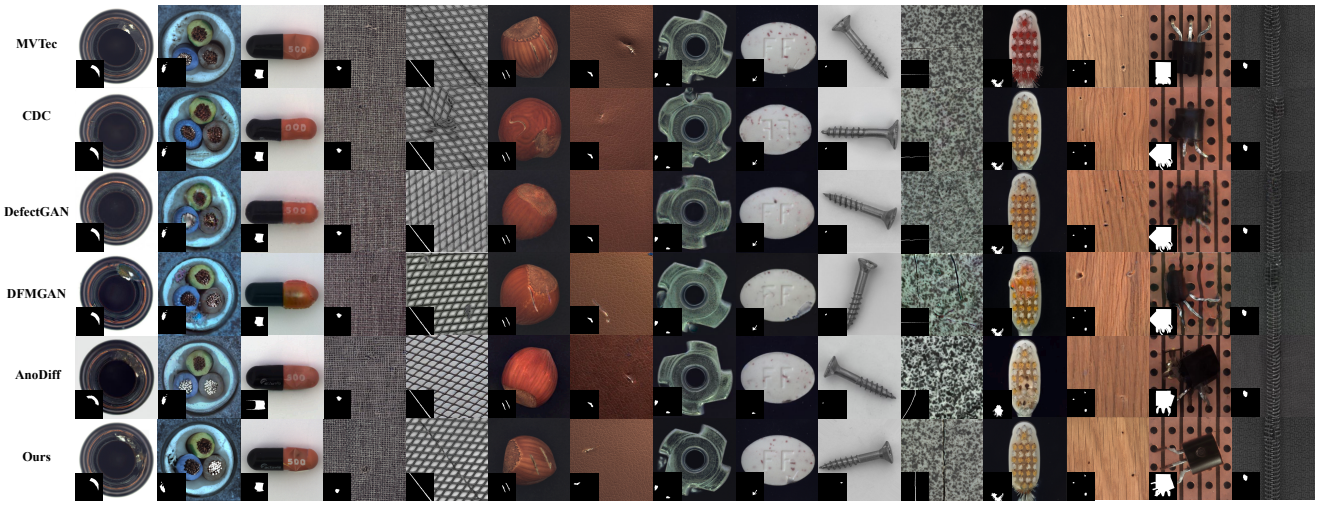


Figure 3. Selected generation qualitative results on MVTec AD.

Method	MVTec AD			VisA		
	i-AUC	p-AUC	Pro	i-AUC	p-AUC	Pro
SimpleNet [7]	99.5	98.1	90.0	97.1	98.2	90.7
GLAD [8]	99.3	98.6	95.2	99.5	98.6	94.3
GLASS [1]	**99.9**	99.3	96.8	98.8	98.8	92.2
Ours + GLASS	**99.9**	**99.4**	**97.2**	**99.5**	**98.9**	**95.3**

Table 7. Quantitative Results on MVTec AD and VisA for **one-to-one** model between our method and SOTA methods.

+0.3% on MVTec AD, and achieves a +1.4% improvement in image-level AP on VisA. Importantly, these improvements are achieved without adding model complexity, as our FLOPs remain at just 43% of those required by MambaAD. These comparisons demonstrate that the images generated by our approach can be effectively applied to lightweight segmentation models, enhancing their performance while maintaining computational efficiency. This highlights the potential of our method in resource-constrained industrial anomaly detection scenarios.

Comparision with generated method. We present the detailed quantitative results of our method compared to representative generative approaches on the anomaly inspection downstream task for each category in the dataset. Tab. 8

Category	DFMGAN			AnoDiff			Ours			Ours(one-shot)		
	AUC	AP	F_1-max	AUC	AP	F_1-max	AUC	AP	F_1-max	AUC	AP	F_1-max
bottle	99.3	99.8	97.7	99.8	**99.9**	**98.9**	**99.9**	**99.9**	**98.9**	99.7	99.8	98.8
cable	95.9	97.8	93.8	**100**	**100**	**100**	99.0	99.1	95.5	98.3	98.4	94.0
capsule	92.8	98.5	94.5	**99.7**	**99.9**	**98.7**	98.7	99.6	97.3	96.4	99.0	94.8
carpet	67.9	87.9	87.3	96.7	98.8	94.3	**98.9**	**99.6**	**98.4**	97.5	99.0	96.7
grid	73.0	90.4	85.4	98.4	99.5	98.7	**100**	**100**	**100**	**100**	**100**	**100**
hazelnut	**99.9**	**100**	**99.9**	99.8	99.9	98.9	**99.9**	99.9	**99.9**	**99.9**	**100**	99.0
leather	99.9	**100**	99.2	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**
metal_nut	99.3	99.8	99.2	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**
pill	68.7	91.7	91.4	98.0	99.6	97.0	**99.2**	**99.8**	**97.9**	98.6	99.6	96.9
screw	22.3	64.7	85.3	96.8	97.9	95.5	**98.8**	**99.4**	**95.8**	90.7	95.2	90.1
tile	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**
toothbrush	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	94.2	96.5	92.3
transistor	90.8	92.5	88.9	**100**	**100**	**100**	**100**	**100**	**100**	98.0	97.1	94.3
wood	98.4	99.4	98.8	98.4	99.4	98.8	**100**	**100**	**100**	99.4	99.7	98.8
zipper	99.7	99.9	99.4	99.9	**100**	99.4	**100**	**100**	**100**	**100**	**100**	**100**
Average	87.2	94.8	94.7	99.2	99.7	98.7	**99.6**	**99.8**	**98.8**	98.2	99.0	97.1

Table 8. **Comparison of anomaly detection on MVTec AD** by training a U-Net on the generated data.

and Tab. 9 report the results on MVTec AD, while Tab. 10 and Tab. 11 present the results on VisA.

2.5. More Qualitative Experiments

Qualitative examples of generation. Fig. 3 and Fig. 4 presents the defect generation results for all objects in MVTec AD and VisA datasets across various methods. DFMGAN [2] struggles to fully capture the appearance of complex objects, often producing highly distorted images

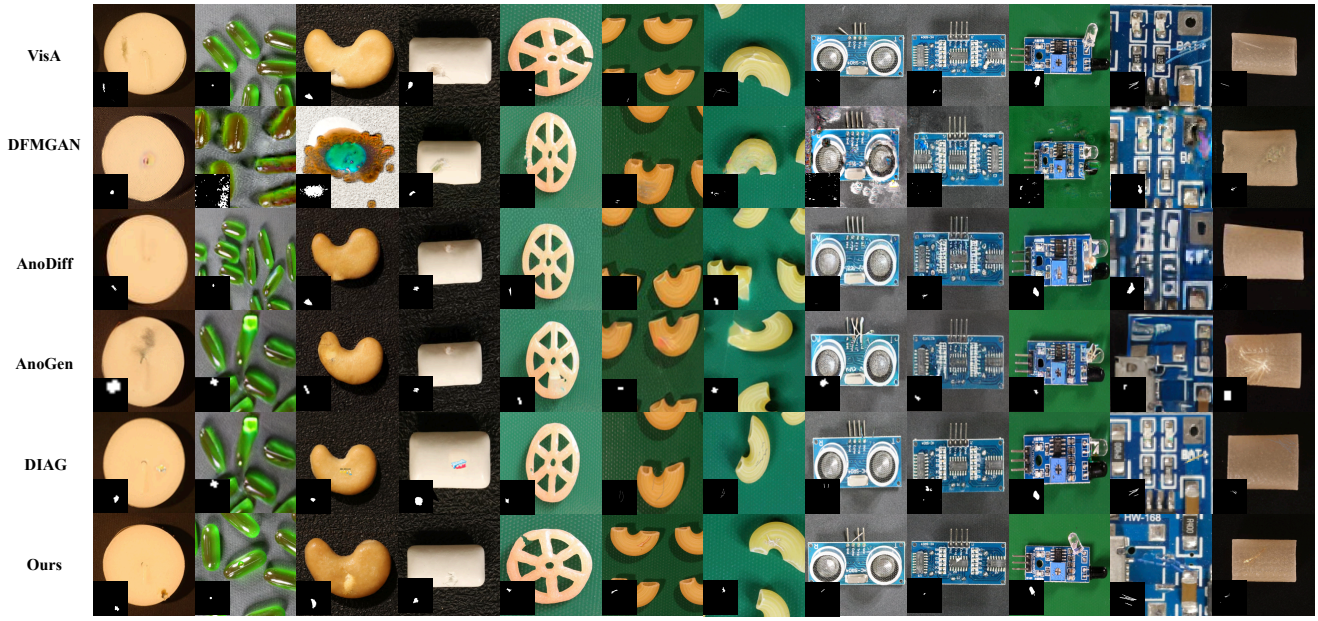


Figure 4. Selected generation qualitative results on VisA.

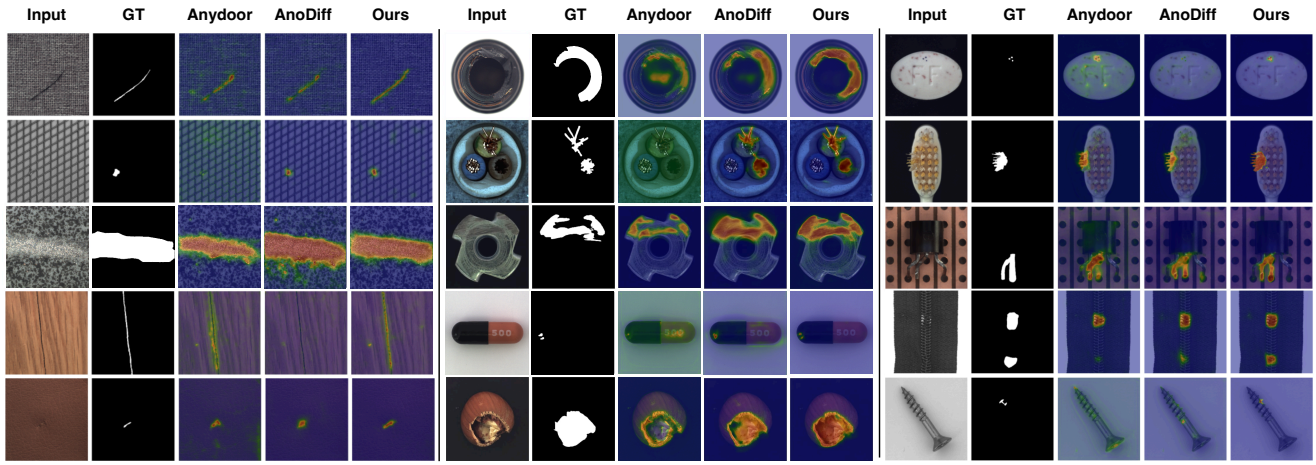


Figure 5. Selected localization qualitative results on MVTec AD.

that fail to preserve structural integrity. Additionally, it cannot guarantee alignment between the generated anomalies and their corresponding masks. The state-of-the-art AnoDiff [6] encounters difficulties when handling small defects and intricate objects, such as the transistors and bulbs on PCBs. AnoGen [3], which generates anomalies using bounding box masks, achieves notable improvements in blending defects with the background. However, it frequently introduces artifacts in the defective regions. Meanwhile, the training-free method DIAG [6], which employs text prompts to generate anomalies, produces defects that significantly deviate from real-world characteristics, failing to align with the semantic nature of industrial anomalies. In contrast, our approach excels in generating precise tiny defects while maintaining accurate mask alignment.

Qualitative examples of localization. Fig. 5 and Fig. 6 illustrates the qualitative results by generating a heatmap for suspicious regions on MVTec AD and VisA datasets. The first two columns exhibit the abnormal images and their corresponding ground truth, while the remaining columns present the predictions of different methods, including Anydoor, AnoDiff, and our TF-IDG. Notably, the detection model trained on images produced by Anydoor struggles to discern defects that closely resemble the background, such as missing wires in cables, cracks in bottles, chunk in candle, or same color spot in cashew. AnoDiff demonstrates low sensitivity to minor defects, such as those found in wood, screws or PCB. In contrast to the abovementioned

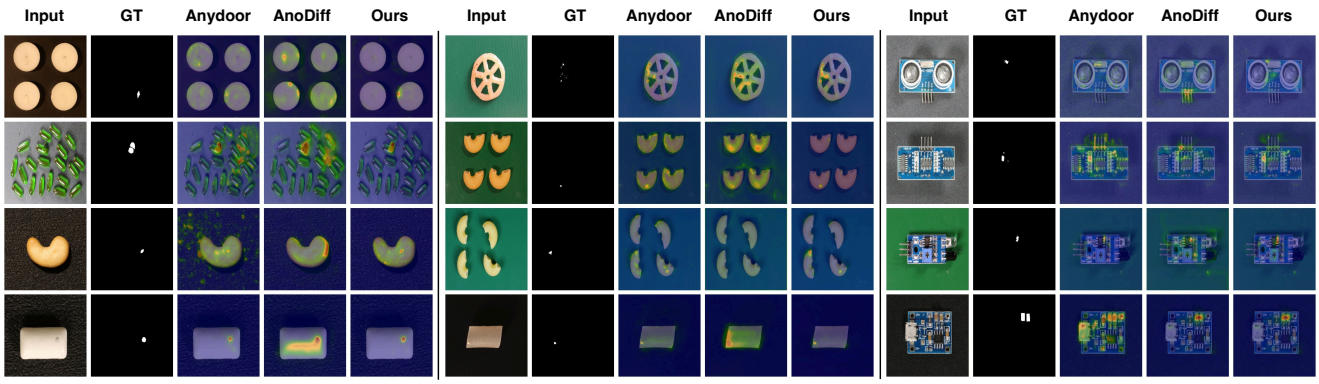


Figure 6. Selected localization qualitative results on VisA.

Category	DFMGAN			AnoDiff			Ours			Ours(one-shot)		
	AUC	AP	F_1-max	AUC	AP	F_1-max	AUC	AP	F_1-max	AUC	AP	F_1-max
bottle	98.9	90.2	83.9	**99.4**	**94.1**	**87.3**	99.3	88.0	79.9	98.0	89.8	83.4
cable	97.2	81.0	75.4	**99.2**	**90.8**	**83.5**	98.1	81.3	75.7	95.8	73.0	68.4
capsule	79.2	26.0	35.0	**98.8**	**57.2**	**59.8**	98.2	53.9	58.2	98.0	57.7	57.5
carpet	90.6	33.4	38.1	98.6	81.2	74.6	**99.0**	86.1	**78.9**	98.7	85.4	78.6
grid	75.2	14.3	20.5	98.3	52.9	54.6	**99.6**	**69.4**	**64.0**	99.5	63.8	61.3
hazelnut	99.7	95.2	89.5	**99.8**	**96.5**	**90.6**	98.8	85.2	79.0	96.5	70.9	71.8
leather	98.5	68.7	66.7	**99.8**	79.6	71.0	**99.8**	82.1	**75.3**	**99.8**	**82.7**	74.4
metal nut	99.3	98.1	**94.5**	**99.8**	**98.7**	94.0	98.9	94.8	88.9	98.9	95.2	88.3
pill	81.2	67.8	72.6	**99.8**	**97.0**	**90.8**	99.5	93.8	86.3	98.8	91.3	85.2
screw	58.8	2.20	5.30	97.0	51.8	50.9	**99.2**	**79.6**	**74.1**	96.2	19.9	27.1
tile	**99.5**	**97.1**	**91.6**	99.2	93.9	86.2	**99.5**	94.6	86.5	99.0	93.1	86.5
toothbrush	96.4	75.9	72.6	99.2	76.5	73.4	**99.8**	**92.8**	**86.0**	92.4	33.2	38.6
transistor	96.2	81.2	77.0	**99.3**	**92.6**	85.7	99.1	89.2	**85.9**	95.3	79.0	75.1
wood	95.3	70.7	65.8	**98.9**	**84.6**	74.5	98.5	82.8	**75.4**	94.8	75.3	70.5
zipper	92.9	65.6	64.9	**99.4**	86.0	79.2	99.2	88.3	**80.9**	99.3	88.3	**80.9**
Average	90.0	62.7	62.1	**99.1**	81.4	76.3	**99.1**	**84.1**	**78.3**	97.4	73.2	69.8

Table 9. **Comparison of anomaly localization on MVTec AD** by training a U-Net on the generated data.

Category	DFMGAN			AnoDiff			Ours			Ours (one-shot)		
	AUC	AP	F_1-max	AUC	AP	F_1-max	AUC	AP	F_1-max	AUC	AP	F_1-max
candle	83.4	15.5	25.4	96.9	28.8	34.5	**98.8**	47.5	46.5	**98.8**	**48.7**	**47.9**
capsules	60.0	1.5	5.3	97.2	56.4	57.9	**99.1**	**74.8**	**71.2**	98.9	69.5	67.7
cashew	88.6	5.0	9.6	**99.4**	**87.9**	**84.3**	98.0	74.8	71.3	96.1	19.0	25.6
chewinggum	98.5	82.4	75.9	98.1	58.5	57.5	99.5	**85.7**	**78.9**	**99.7**	85.1	77.2
fryum	93.0	23.1	29.5	**98.4**	**65.7**	**64.1**	96.3	47.4	49.5	95.2	39.2	42.8
macaroni1	93.4	21.5	30.3	94.1	5.4	15.3	**99.5**	**49.2**	**52.6**	99.4	28.5	38.1
macaroni2	94.2	1.6	3.9	92.3	0.1	0.3	**99.1**	**37.9**	**42.5**	98.9	30.7	35.7
pcb1	84.4	34.9	40.3	**96.0**	**68.1**	**70.0**	92.0	67.1	69.9	94.9	37.9	42.4
pcb2	91.6	27.0	37.4	95.2	21.5	29.0	95.8	**54.7**	**56.1**	**97.5**	42.4	51.6
pcb3	81.9	7.8	16.8	**97.5**	27.9	33.6	95.3	**57.7**	**58.3**	92.0	45.5	47.8
pcb4	95.2	44.5	47.2	**98.9**	**60.2**	**56.9**	98.2	41.5	43.0	96.7	35.0	43.4
pipe fryum	90.5	17.8	25.1	**99.5**	**76.7**	**69.4**	99.0	73.9	67.0	97.6	50.6	49.3
Average	87.9	23.6	28.9	97.0	46.4	47.7	**97.6**	**59.4**	**58.9**	97.1	44.3	47.4

Table 10. **Comparison of anomaly localization on VisA** by training a U-Net on the generated data.

methods, our TF-IDG precisely identifies anomalies and mitigates the false detection in the background.

Category	DFMGAN			AnoDiff			Ours			Ours (one-shot)		
	AUC	AP	F_1-max	AUC	AP	F_1-max	AUC	AP	F_1-max	AUC	AP	F_1-max
candle	86.8	84.3	75.6	95.4	94.8	88.5	**97.4**	**96.3**	**90.5**	96.0	94.5	87.2
capsules	72.8	75.1	74.0	90.2	92.7	85.5	**98.5**	**98.7**	**94.2**	96.1	97.0	90.6
cashew	85.7	87.0	84.1	89.7	91.9	85.9	**97.5**	**98.2**	**93.7**	93.7	94.8	91.4
chewinggum	98.8	99.2	96.2	87.9	92.8	84.1	97.7	98.8	96.2	**99.2**	**99.5**	**97.1**
fryum	80.3	87.2	78.8	83.6	89.0	80.9	**98.2**	**98.8**	**94.1**	93.0	95.9	88.4
macaroni1	90.8	90.9	82.7	94.1	90.0	87.5	**99.6**	**99.6**	**97.8**	95.5	94.2	89.4
macaroni2	66.1	61.6	61.3	59.3	48.0	58.9	**93.3**	**92.9**	**86.1**	84.3	82.1	75.0
pcb1	90.8	87.9	82.3	**94.5**	**93.3**	**87.0**	94.1	92.6	83.5	87.7	85.8	78.9
pcb2	96.4	95.9	90.2	93.0	94.4	90.1	**99.6**	**99.5**	**97.7**	97.9	97.2	91.3
pcb3	74.3	72.8	63.2	91.6	90.8	81.3	**97.2**	**97.5**	**94.0**	95.3	93.3	90.2
pcb4	97.9	96.9	91.5	98.2	98.0	**95.7**	**98.7**	**98.5**	95.6	98.5	97.8	95.0
pipe fryum	83.0	89.1	80.3	86.4	90.3	85.1	**97.1**	**98.3**	**94.2**	86.8	91.2	84.7
Average	85.3	85.7	80.0	88.7	88.8	84.2	**97.4**	**97.5**	**90.7**	93.7	93.6	88.3

Table 11. **Comparison of anomaly detection on VisA** by training a U-Net on the generated data.

3. Limitation

While our training-free framework demonstrates strong performance in generating localized and texture-level anomalies through image inpainting, we observe limitations in handling semantic or logical defects. This suggests that addressing high-level structural or contextual defects may require incorporating richer object semantics and global contextual priors. Enhancing the model's capacity for such global reasoning represents a promising direction for future research in industrial anomaly synthesis.

References

[1] Qiyu Chen, Huiyuan Luo, Chengkan Lv, and Zhengtao Zhang. A unified anomaly synthesis strategy with gradient ascent for industrial anomaly detection and localization. In *ECCV*, pages 37–54. Springer, 2024. 2, 3

[2] Yuxuan Duan, Yan Hong, Li Niu, and Liqing Zhang. Few-shot defect image generation via defect-aware feature manipulation. In *AAAI*, pages 571–578, 2023. 3

[3] Guan Gui, Bin-Bin Gao, Jun Liu, Chengjie Wang, and Yunsheng Wu. Few-shot anomaly-driven generation for anomaly classification and segmentation. In *ECCV*, pages 210–226. Springer, 2024. 4

[4] Haoyang He, Jiangning Zhang, Hongxu Chen, Xuhai Chen, Zhishan Li, Xu Chen, Yabiao Wang, Chengjie Wang, and Lei

Xie. A diffusion-based framework for multi-class anomaly detection. In *AAAI*, pages 8472–8480, 2024. 2, 3

[5] Haoyang He, Yuhu Bai, Jiangning Zhang, Qingdong He, Hongxu Chen, Zhenye Gan, Chengjie Wang, Xiangtai Li, Guanzhong Tian, and Lei Xie. Mambaad: Exploring state space models for multi-class unsupervised anomaly detection. *NeurIPS*, 37:71162–71187, 2025. 2, 3

[6] Teng Hu, Jiangning Zhang, Ran Yi, Yuzhen Du, Xu Chen, Liang Liu, Yabiao Wang, and Chengjie Wang. Anomalydiffusion: Few-shot anomaly image generation with diffusion model. In *AAAI*, pages 8526–8534, 2024. 4

[7] Zhikang Liu, Yiming Zhou, Yuansheng Xu, and Zilei Wang. Simplenet: A simple network for image anomaly detection and localization. In *CVPR*, pages 20402–20411, 2023. 2, 3

[8] Hang Yao, Ming Liu, Zhicun Yin, Zifei Yan, Xiaopeng Hong, and Wangmeng Zuo. Glad: towards better reconstruction with global and local adaptive diffusion models for unsupervised anomaly detection. In *ECCV*, pages 1–17. Springer, 2024. 2, 3

[9] Zhiyuan You, Lei Cui, Yujun Shen, Kai Yang, Xin Lu, Yu Zheng, and Xinyi Le. A unified model for multi-class anomaly detection. *NeurIPS*, 35:4571–4584, 2022. 2, 3

[10] Xuan Zhang, Shiyu Li, Xi Li, Ping Huang, Jiulong Shan, and Ting Chen. Destseg: Segmentation guided denoising student-teacher for anomaly detection. In *CVPR*, pages 3914–3923, 2023. 2, 3