Wave-MambaAD: Wavelet-driven State Space Model for Multi-class Unsupervised Anomaly Detection
Supplementary Material

Qiao Zhang Mingwen Shao* Xinyuan Chen Xiang Lv Kai Xu
Shandong Key Laboratory of Intelligent Oil & Gas Industrial Software, Qingdao Institute of Software,
College of Computer Science and Technology, China University of Petroleum (East China)
15263653689@163.com, smw278@126.com, xinyua_sdnu@163.com, lvxiang1997@126.com, s23070014@s.upc.edu.cn

Overview

The supplementary material presents the following sections to strengthen the main manuscript:

Sec. A introduces the formulas for the Wavelet Transform and its advantages over the Fourier Transform.

Sec. B introduces the datasets and metrics.

Sec. C shows more ablation studies.

Sec. D shows more quantitative comparisons.

A. Wavelet Transform's Formulas and Advantages

The wavelet transform is a signal analysis method that provides both spatial and frequency domain information about a signal. Through wavelet decomposition, input features can be divided into low-frequency components and high-frequency components. Low-frequency components primarily reflect overall contours and global structural information, while high-frequency components more accurately reflect edges, textures, or subtle structural changes.

Industrial anomalies typically manifest as complex features across multiple scales. Large-scale defects often alter the overall shape or structure of an object, while subtle defects are more concealed within local textures. To address this, we combine the low-frequency components in the wavelet domain with large-scale defect modelling, leveraging high-frequency components to enhance sensitivity to local details, thereby achieving efficient anomaly detection across multiple scales.

To enable multi-scale decomposition and frequency-domain analysis, we employ the Haar discrete wavelet transform (DWT), which factorizes an input feature map $F_{in} \in \mathbb{R}^{C\times H\times W}$ into low-frequency (F_l) and high-frequency (F_h) components as:

$$F_l, F_h^{(LH)}, F_h^{(HL)}, F_h^{(HH)} = \mathbf{DWT}(F_{in}) \tag{1}$$

*Corresponding author

where F_l captures the global structure, while $F_h^{(LH)}$, $F_h^{(HL)}$, $F_h^{(HH)}$ represent horizontal, vertical, and diagonal details, respectively. The DWT utilizes simple orthogonal filters:

$$L = \frac{1}{\sqrt{2}}[1,1], H = \frac{1}{\sqrt{2}}[1,-1] \tag{2}$$

where L and H are the low- and high-pass filters. The inverse wavelet transform (IWT) reconstructs the original feature by:

$$\begin{aligned} F_{out} = \mathbf{IWT}(F_h^{out}, F_l^{out}) = (F_l^{out} * L^T) * L \\ + (F_h^{(LH)} * L^T) * H \\ + (F_h^{(HL)} * H^T) * L \\ + (F_h^{(HH)} * H^T) * H \end{aligned} \tag{3}$$

where $*$ denotes the convolution operation applied along spatial dimensions. Compared with pure spatial-domain processing, this decomposition enables the model to explicitly manipulate global structures and fine-grained details separately, which is beneficial for anomaly detection tasks.

Why Wavelet Transform over Fourier Transform? The wavelet transform decomposes features into low-frequency and directionally-aware high-frequency components, enabling fine-grained modeling of subtle defects and adaptive preservation of global structures. Unlike the Fourier transform, which captures only global frequency information without spatial localization, the wavelet transform provides joint spatial-frequency analysis. This is particularly beneficial for anomaly detection, as defects often exhibit localized, orientation-dependent characteristics. By capturing both spatial position and directional frequency details, wavelet-based methods effectively detect subtle anomalies while accurately localizing large-scale structural defects. In contrast, Fourier-based approaches may overlook such irregularities due to their inherently global, non-localized nature.

B. Datasets and Metrics

The comparative datasets and metrics are described below: **MVTec-AD**[1] includes 5,354 high-resolution images from 5 texture and 10 object categories, with 3,629 normal images for training and 1,725 for testing. **VisA**[10] contains 10,821 images of 12 objects, with 9,621 normal samples and 1,200 anomalies, covering surface and structural defects. **MPDD** [6] is designed for inspecting metal part defects, containing 1,346 images of 6 metal products. **MVTec-3D** [2] is a 3D dataset for unsupervised anomaly detection and localisation, containing over 4,000 high-resolution scans across 10 object categories with both defect-free and defective samples. **Real-IAD** [9] is a large-scale real-world industrial anomaly detection dataset consisting of 150,000 high-resolution images covering 30 different objects.

AU-ROC measures the model's ability to distinguish normal and abnormal samples across thresholds. **AP** evaluates the precision-recall balance, with higher values indicating better anomaly detection. **F1_max** assesses the trade-off between precision and recall at various thresholds. **AU-PRO** quantifies the precision-recall trade-off, especially for imbalanced data, reflecting anomaly localization performance. AU-ROC, AP, and F1_max are evaluated at both pixel and image levels.

C. Ablation Studies

We present a series of ablation experiments on the MVTec-AD dataset in Tab. 1, Tab. 2, and Tab. 3, further to validate the effectiveness of the proposed method setup.

Effectiveness of setting the number of Wavelet-Mamba modules per layer in the Wavelet-Mamba decoder. We show the ablation experiments on the number of Wavelet-Mamba modules per layer in the Wavelet-Mamba decoder in Tab. 1. Specifically, we use the number of Wavelet-Mamba modules in the decoder set from shallow to deep [1, 1, 1, 1] as a baseline and then increase the number from deep to shallow. We find the best results with the setting of [1, 2, 2, 2] with acceptable parameters and complexity.

Verification of the effectiveness of directional selective scanning in HFSS for high-frequency components in different directions. Tab. 2 shows ablation experiments replacing directional selective scanning with different scanning strategies. We find that directional selection scanning achieves better performance, as it aligns with high-frequency subbands and can capture subtle defects in the corresponding direction.

The validity of the channel group number setting in LFSS. Tab. 3 demonstrates the ablation experiments for the number of channel grouping settings in LFSS. The best results are achieved when the number of channel grouping settings in LFSS in each layer of the Wavelet-Mamba module in the decoder is set to [2, 4, 8, 16] from shallow to deep. Additionally, we observed that improper channel grouping leads to significant performance degradation. Over-grouping of shallow-layer features causes channel fragmentation, while insufficient grouping of deep-layer features reduces the expression of high-level semantic information. The pyramid-style channel grouping effectively avoids over-segmentation of shallow-layer features by adjusting the number of groups based on channel depth, thereby enhancing the understanding of deep-layer semantic information.

D. Quantitative Comparison

Performance comparison with MambaAD under different Seeds. To address concerns regarding performance stability, we conducted experiments with five different random seeds for both MambaAD and our proposed Wave-MambaAD. Results show that Wave-MambaAD achieves an average performance of 85.58 ± 0.14, consistently outperforming MambaAD (85.40 ± 0.13) with stable improvements across runs (see Table 4).

Quantitative comparison with other SOTA on MVTec-3D [2]. Tab. 5 shows a quantitative comparison between the proposed method and SOTA on MVTec-3D. Our method is weaker than DiAD in metric F1_max, but achieves the best performance in other metrics, further demonstrating the effectiveness of our method.

Quantitative comparison with other SOTA on Real-IAD [9]. Tab. 6 shows a quantitative comparison between the proposed method and SOTA on Real-IAD. Our method achieved the best performance across all metrics, further validating the robustness of the proposed method in real industrial scenarios.

Quantitative comparison of categories on the MVTec-AD [1]. Tab. 7 and Tab. 8 present the image-level anomaly detection results and pixel-level anomaly localization quantitative results for all categories in the MVTec-AD dataset, respectively. The results demonstrate that the performance of our method is comparable to other SoTA methods.

Quantitative comparison of categories on the Visa [10]. Tab. 9 and Tab. 10 present the image-level anomaly detection results and pixel-level anomaly localization quantitative results for all categories in the Visa dataset, respectively. Our method can achieve the best performance at pixel-level AU-ROC and competitive performance at other pixel or image levels.

Quantitative comparison of categories on the MPDD [6]. Tab. 11 and Tab. 12 present the image-level anomaly detection results and pixel-level anomaly localization quantitative results for all categories in the MPDD industrial dataset, respectively. Specifically, our method outperforms the baseline in quantitative comparisons at the pixel level, achieving the best performance on the categories' bracket_black', 'connector', 'metal_plate', and 'tubes'. Furthermore, our method also achieves competitive performance in quantitative comparisons at the image level.

Quantitative comparison of categories on the MVTec-3D [2]. Tab. 13 and Tab. 14 show the image-level anomaly detection results and pixel-level anomaly localisation quantitative results for all categories in the MVTec-3D industrial dataset, respectively. Specifically, our method outperforms other SOTA methods in pixel-level average quantitative comparisons, and it is also competitive in image-level average quantitative comparisons.

Quantitative comparison of categories on the Real-IAD [9]. Tab. 15 and Tab. 16 show the image-level anomaly detection results and pixel-level anomaly localisation quantitative results for all categories in the Real-IAD industrial dataset, respectively. Specifically, our method achieves the best performance in most categories, far surpassing other SOTA in categories such as 'fire hood', "regulator" and 'toy brick'.

References

[1] Paul Bergmann, Michael Fauser, David Sattlegger, and Carsten Steger. Mvtec ad–a comprehensive real-world dataset for unsupervised anomaly detection. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 9592–9600, 2019. 2

[2] Paul Bergmann, Xin Jin, David Sattlegger, and Carsten Steger. The mvtec 3d-ad dataset for unsupervised 3d anomaly detection and localization. *arXiv preprint arXiv:2112.09045*, 2021. 2, 3, 4

[3] Denis Gudovskiy, Shun Ishizaka, and Kazuki Kozuka. Cflow-ad: Real-time unsupervised anomaly detection with localization via conditional normalizing flows. In *Proceedings of the IEEE/CVF winter conference on applications of computer vision*, pages 98–107, 2022. 4

[4] 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. *arXiv preprint arXiv:2404.06564*, 2024. 4

[5] 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 *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 8472–8480, 2024. 4

[6] Stepan Jezek, Martin Jonak, Radim Burget, Pavel Dvorak, and Milos Skotak. Deep learning-based defect detection of metal parts: evaluating current methods in complex conditions. In *2021 13th International congress on ultra modern telecommunications and control systems and workshops (ICUMT)*, pages 66–71. IEEE, 2021. 2

[7] Jiarui Lei, Xiaobo Hu, Yue Wang, and Dong Liu. Pyramidflow: High-resolution defect contrastive localization using pyramid normalizing flow. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 14143–14152, 2023. 4

[8] 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. 4

[9] 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 *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 22883–22892, 2024. 2, 3, 4

[10] Yang Zou, Jongheon Jeong, Latha Pemula, Dongqing Zhang, and Onkar Dabeer. Spot-the-difference self-supervised pretraining for anomaly detection and segmentation. In *European Conference on Computer Vision*, pages 392–408. Springer, 2022. 2

Settings	Image-level AU-ROC	AP	F1_max	Pixel-level AU-ROC	AP	F1_max	AU-PRO	Params(M)	FLOPs(G)
$[1,1,1,1]$	97.2	98.9	96.9	97	51.5	55.7	91.6	**17.2**	**5.3**
$[1,1,1,2]$	97.6	99.0	97.0	97.0	53.5	56.9	92.3	17.5	6.1
$[1,1,2,2]$	98.0	99.1	97.1	97.4	54.6	57.6	92.9	18.6	6.8
$[1,2,2,2]$	**98.2**	**99.3**	**97.4**	**97.5**	**55.9**	**58.3**	**93.1**	22.3	7.5
$[2,2,2,2]$	98.1	98.9	97.4	97.2	54.8	57.1	92.3	37.8	8.2

Table 1. **Ablation studies on the number of Wavelet-Mamba Modules per layer setting in the Wavelet-Mamba Decoder.** The best results are shown in **bold**.

Setting	Image-level AU-ROC	AP	F1_max	Pixel-level AU-ROC	AP	F1_max	AU-PRO
HFSS(Scan)	98.0	99.2	97.1	97.4	55.0	57.7	92.5
HFSS(Zorder)	98.1	**99.3**	97.2	97.3	54.4	57.5	92.3
HFSS(Zigzag)	**98.2**	99.2	97.3	97.0	54.4	57.7	92.3
HFSS(Hilbert)	98.1	**99.3**	97.1	97.2	54.6	57.7	92.6
HFSS(DSS)	**98.2**	**99.3**	**97.4**	**97.5**	**55.9**	**58.3**	**93.1**

Table 2. **Ablation experiments with different scanning in HFSS.** The best results are shown in **bold**.

Settings	Image-level AU-ROC	AP	F1_max	Pixel-level AU-ROC	AP	F1_max	AU-PRO
$[1,1,1,1]$	93.6	97.6	95.4	96.7	49.1	53.5	90.5
$[2,2,2,2]$	98.0	99.2	97.2	97.1	54.2	57.6	92.3
$[4,4,4,4]$	98.0	**99.3**	97.3	97.1	54.5	57.5	92.5
$[8,8,8,8]$	97.8	99.1	96.8	97.0	53.3	56.9	92.1
$[2,4,8,16]$	**98.2**	**99.3**	**97.4**	**97.5**	**55.9**	**58.3**	**93.1**

Table 3. **Ablation experiments on the number of channel groupings in LFSS.** The best results are shown in **bold**.

Method	Seed=21	Seed=42	Seed=63	Seed=84	Seed=126
MambaAD	85.4	85.4	85.6	85.2	85.4
Wave-MambaAD	85.4	85.6	85.8	85.5	85.6

Table 4. **Performance of the proposed method with MambaAD under different Seed settings.**

Datasets	Method	Image-level AU-ROC	AP	F1_max	Pixel-level AU-ROC	AP	F1_max	AU-PRO	mAD
MVTec-3D [2]	CFLOW-AD [3]	73.1	91.0	90.2	<u>96.8</u>	21.6	26.6	89.0	69.8
	SimpleNet [8]	75.8	92.3	90.4	<u>94.7</u>	17.3	23.4	81.0	67.8
	PyramidalFlow [7]	58.7	85.6	88.6	90.8	7.2	12.0	74.9	59.7
	DiAD [5]	<u>84.6</u>	94.8	**95.6**	96.4	25.3	32.3	87.8	73.8
	MambaAD [4]	<u>84.1</u>	<u>95.1</u>	<u>92.2</u>	**98.6**	<u>36.9</u>	<u>40.8</u>	<u>94.2</u>	<u>77.4</u>
	Wave-MambaAD	**85.1**	**95.5**	<u>92.2</u>	**98.6**	**37.4**	**41.6**	**94.4**	**77.8**

Table 5. **Quantitative comparison on the MVTec-3D dataset.** The best results are denoted using **bold** and the second-best results are <u>underlined</u>.

Datasets	Method	Image-level AU-ROC	AP	F1_max	Pixel-level AU-ROC	AP	F1_max	AU-PRO	mAD
Real-IAD	CFLOW-AD [3]	77.0	75.8	69.9	<u>94.8</u>	17.6	21.7	80.4	62.5
	SimpleNet [8]	57.2	53.4	61.5	<u>76.1</u>	1.9	4.9	42.4	42.5
	PyramidalFlow [7]	54.5	47.9	62.0	71.1	1.2	1.1	35.8	39.1
	DiAD [5]	75.6	66.4	69.9	88.0	2.9	7.1	58.0	52.6
	MambaAD [4]	<u>87.0</u>	<u>85.3</u>	<u>77.6</u>	**98.6**	<u>32.4</u>	<u>38.1</u>	<u>91.2</u>	<u>72.9</u>
	Wave-MambaAD	**88.9**	**87.2**	**79.6**	**98.6**	**37.7**	**42.5**	**91.7**	**75.2**

Table 6. **Quantitative comparison on the Real-IAD [9] dataset.** The best results are denoted using **bold** and the second-best results are <u>underlined</u>.

Method	CFLOW-AD			SimpleNet			PyramidalFlow			RealNet			DiAD			MambaAD			Wave-MambaAD		
Category	WACV22			CVPR23			CVPR23			CVPR24			AAAI24			NeurIPS24			Ours		
candle	93.0	93.3	85.3	93.6	94.5	84.9	43.4	44.9	66.7	52.7	57.8	66.7	92.8	92	87.6	96.8	96.9	90.1	95.9	96	**90.6**
capsules	54.8	72.8	76.9	76.8	87.0	78.3	55.8	69.4	76.9	72.8	85.0	77.0	58.2	69.0	78.5	**91.8**	**95.0**	**88.8**	90.7	94.0	88.6
cashew	95.8	98.2	93.3	93.2	96.8	90.2	84.0	90.7	87.7	79.8	89.6	84.5	91.5	95.7	89.7	**94.5**	**97.3**	**91.1**	93.0	96.7	90.7
chewinggum	97.3	98.9	95.3	97.4	98.8	93.8	36.3	63.6	80.0	87.7	94.6	86.0	**99.1**	**99.5**	**95.9**	97.7	98.9	94.2	97.3	98.8	94.7
fryum	80.1	95.5	88.3	84.0	92.8	83.6	68.7	83.5	81.0	66.2	83.4	80.6	89.8	95.0	87.2	95.2	97.7	90.5	**95.6**	**97.9**	**91.5**
macaroni1	80.6	77.5	75.3	77.4	75.7	71.4	41.9	43.4	66.0	76.2	81.5	72.0	85.7	85.2	78.8	91.6	**89.8**	81.6	**92.0**	**89.8**	**83.6**
macaroni2	64.3	62.7	66.2	66.7	61.2	66.9	67.8	71.0	69.2	57.6	57.5	69.4	62.5	57.4	69.6	81.6	78.0	73.8	**87.5**	**85.3**	**79.6**
pcb1	93.5	93.5	86.5	92.8	93.9	87.9	83.3	80.9	82.1	76.6	79.6	71.2	88.1	88.7	80.7	**95.4**	**93.0**	91.6	95.0	92.9	**92.3**
pcb2	92.4	93	85.8	90.1	91.8	83.4	72.8	76.3	70.8	71.6	78.8	67.4	91.4	91.4	84.7	94.2	**93.7**	89.3	**94.7**	**93.7**	**89.6**
pcb3	80.5	83.6	73.6	86.4	88.3	78.0	53.8	56.4	66.4	78.8	84.0	72.5	86.2	87.6	77.6	**93.7**	**94.1**	**86.7**	93.4	93.8	86.3
pcb4	98.5	98.4	96.1	97.6	97.8	92.0	48.9	52.7	66.4	72.5	79.2	68.3	99.6	99.5	97.0	**99.9**	**99.9**	**98.5**	**99.9**	**99.9**	98.0
pipe_fryum	97.4	98.8	96.0	80.7	90.9	83.2	41.6	63.3	80.0	64.0	83.5	80.5	96.2	98.1	93.7	**98.7**	**99.3**	**97.3**	98.4	99.2	96.0
Average	86.5	88.8	84.9	86.4	89.1	82.8	58.2	66.3	74.4	71.4	79.5	74.7	86.8	88.3	85.1	94.3	94.5	89.4	**94.4**	**94.8**	**90.1**

Table 7. **Comparison of image-level AU-ROC/AP/F1_max metrics with SoTA on the MVTec-AD dataset.** The best results are shown in **bold**.

Method	CFLOW-AD				SimpleNet				PyramidalFlow				RealNet				DiAD				MambaAD				Wave-MambaAD			
Category	WACV22				CVPR23				CVPR23				CVPR24				AAAI24				NeurIPS24				Ours			
Bottle	97.4	61.8	63.6	92.1	97.3	57.3	65.9	89.1	77.7	16.9	23.9	40.9	69.8	53.9	46.6	60.9	98.4	52.2	54.8	86.6	**98.7**	**79.7**	**75.7**	**96.0**	**98.7**	77.8	75.2	95.6
Cable	89.8	27.3	33.3	78.0	96.6	48.5	55.0	86.0	82.8	19.0	16.4	41.6	61.5	23.5	25.9	33.3	**96.8**	**50.1**	**57.8**	80.5	95.2	42.0	47.9	90.4	96.6	45.9	51.0	**91.3**
Capsule	98.5	41.0	44.2	92.9	98.1	36.9	46.8	87.1	90.3	13.5	19.6	57.3	54.6	23.7	12.1	23.4	97.1	42.0	45.3	87.2	98.2	43.5	**47.7**	**93.0**	98.4	42.3	45.8	**93.0**
Carpet	98.8	55.9	59.4	94.3	97.7	42.8	46.9	90.0	79.2	8.5	15.1	52.3	89.2	69.0	64.3	84.0	98.6	**42.2**	**46.4**	90.6	**99.2**	64.0	63.8	**97.3**	**99.2**	64.0	63.4	96.8
Grid	92.9	18.2	25.1	81.0	96.9	26.4	33.0	88.3	85.7	9.8	16.6	66.9	82.6	41.2	45.7	77.7	96.6	**66.0**	**64.1**	94.0	98.9	48.4	48.5	96.9	**99.2**	47.9	48.1	**97.2**
Hazelnut	98.5	59.1	57.9	95.6	98.1	49.0	52.4	93.9	92.7	33.2	36.9	84.2	77.5	44.2	48.9	75.4	98.3	**79.2**	80.4	91.5	**99.1**	67.0	**66.1**	95.3	98.8	61.1	61.6	**95.4**
Leather	99.2	45.0	46.1	98.1	98.5	27.8	33.9	95.5	87.7	6.4	15.2	74.0	97.9	70.4	68.0	98.0	98.8	**56.1**	**62.3**	91.3	99.3	50.6	50.4	**98.7**	**99.4**	51.3	51.5	98.5
Matel_nut	96.0	71.2	71.7	88.5	97.7	82.2	79.2	87.6	81.6	41.9	45.8	37.3	52.5	32.3	21.0	39.6	**97.3**	30.0	38.3	90.6	96.7	**74.2**	**78.3**	**92.9**	96.8	72.8	77.9	**92.9**
Pill	96.7	59.5	56.3	90.5	96.7	74.7	70.6	85.3	83.3	18.3	26.0	65.2	54.4	47.8	8.9	35.0	95.7	46.0	51.4	89.0	96.2	55.0	58.9	95.2	**97.5**	**65.5**	**67.5**	**96.1**
Screw	96.5	13.6	18.5	87.7	95.8	15.8	23.8	83.1	71.4	0.9	2.1	21.5	51.8	15.4	4.5	18.5	97.9	**60.6**	**59.6**	95	99.3	45.2	45.1	**97.0**	**99.4**	49.4	49.9	96.8
Tile	96.0	56.0	62.1	86.5	95.4	59.1	60.4	82.5	75.7	55.1	28.1	34.4	93.9	84.1	76.8	90.5	92.4	**65.7**	64.1	**90.7**	93.0	43.9	52.6	79.5	**93.7**	45.4	54.9	81.3
Toothbrush	98.2	45.7	47.2	84.5	98.0	53.6	55.9	80.6	73.0	42.3	31.3	23.2	84.8	50.1	56.1	34.1	**99.0**	**78.7**	**72.8**	**95**	98.9	47.5	59.7	92.0	**99.0**	49.3	60.3	92.0
Transistor	84.8	38.3	39.0	73.0	95.4	60.4	57.6	82.5	75.9	13.2	19.5	26.1	60.9	40.2	28.3	44.6	95.1	15.6	**31.7**	90.0	**96.0**	**63.8**	61.6	**90.3**	93.6	59.0	58.5	85.0
Wood	94.2	45.7	49.8	90.2	92.5	40.3	42.4	80.0	62.6	39.5	17.9	32.3	90.4	76.1	71.9	88.8	93.3	43.3	43.5	**97.5**	94.0	**46.9**	48.4	92.0	**94.5**	**46.9**	**49.1**	90.4
Zipper	97.9	50.2	54.8	92.4	97.9	57.8	55.4	91.9	81.0	15.4	16.3	55.7	67.6	51.9	42.6	47.7	96.2	**60.7**	60.0	91.6	98.1	55.0	58.9	**94.5**	**98.3**	59.3	**60.3**	94.3
Average	95.7	45.9	48.6	88.3	96.8	48.8	51.9	86.9	80.0	22.3	22.0	47.5	72.6	48.3	41.4	56.8	96.8	52.6	55.5	90.7	97.4	55.1	57.6	**93.4**	**97.5**	**55.9**	**58.3**	93.1

Table 8. **Comparison of Pixel-level AU-ROC/AP/F1_max/AU-PRO metrics with SoTA on the MVTec-AD dataset.** The best results are shown in **bold**.

Method	CFLOW-AD			SimpleNet			PyramidalFlow			RealNet			DiAD			MambaAD			Wave-MambaAD		
Category	WACV22			CVPR23			CVPR23			CVPR24			AAAI24			NeurIPS24			Ours		
candle	93.0	93.3	85.3	93.6	94.5	84.9	43.4	44.9	66.7	52.7	57.8	66.7	92.8	92.0	87.6	96.8	96.9	90.1	95.9	96.0	**90.6**
capsules	54.8	72.8	76.9	76.8	87.0	78.3	55.8	69.4	76.9	72.8	85.0	77.0	58.2	69.0	78.5	**91.8**	**95**	**88.8**	90.7	94.0	88.6
cashew	95.8	98.2	93.3	93.2	96.8	90.2	84.0	90.7	87.7	79.8	89.6	84.5	91.5	95.7	89.7	**94.5**	**97.3**	**91.1**	93.0	96.7	90.7
chewinggum	97.3	98.9	95.3	97.4	98.8	93.8	36.3	63.6	80.0	87.7	94.6	86.0	**99.1**	**99.5**	**95.9**	97.7	98.9	94.2	97.3	98.8	94.7
fryum	80.1	95.5	88.3	84	92.8	83.6	68.7	83.5	81.0	66.2	83.4	80.6	89.8	95.0	87.2	95.2	97.7	90.5	**95.6**	**97.9**	**91.5**
macaroni1	80.6	77.5	75.3	77.4	75.7	71.4	41.9	43.4	66.0	76.2	81.5	72.0	85.7	85.2	78.8	91.6	**89.8**	81.6	**92.0**	**89.8**	**83.6**
macaroni2	64.3	62.7	66.2	66.7	61.2	66.9	67.8	71.0	69.2	57.6	57.5	69.4	62.5	57.4	69.6	81.6	78.0	73.8	**87.5**	**85.3**	**79.6**
pcb1	93.5	93.5	86.5	92.8	93.9	87.9	83.3	80.9	82.1	76.6	79.6	71.2	88.1	88.7	80.7	**95.4**	**93.0**	91.6	95.0	92.9	**92.3**
pcb2	92.4	93.0	85.8	90.1	91.8	83.4	72.8	76.3	70.8	71.6	78.8	67.4	91.4	91.4	84.7	94.2	**93.7**	89.3	**94.7**	**93.7**	**89.6**
pcb3	80.5	83.6	73.6	86.4	88.3	78.0	53.8	56.4	66.4	78.8	84	72.5	86.2	87.6	77.6	**93.7**	**94.1**	**86.7**	93.4	93.8	86.3
pcb4	98.5	98.4	96.1	97.6	97.8	92.0	48.9	52.7	66.4	72.5	79.2	68.3	99.6	99.5	97.0	**99.9**	**99.9**	**98.5**	**99.9**	**99.9**	98.0
pipe_fryum	97.4	98.8	96.0	80.7	90.9	83.2	41.6	63.3	80.0	64.0	83.5	80.5	96.2	98.1	93.7	**98.7**	**99.3**	**97.3**	98.4	99.2	96.0
Average	86.5	88.8	84.9	86.4	89.1	82.8	58.2	66.3	74.4	71.4	79.5	74.7	86.8	88.3	85.1	94.3	94.5	89.4	**94.4**	**94.8**	**90.1**

Table 9. **Comparison of image-level AU-ROC/AP/F1_max metrics with SoTA on the Visa dataset.** The best results are shown in **bold**.

Method	CFLOW-AD				SimpleNet				PyramidalFlow				RealNet				DiAD				MambaAD				Wave-MambaAD			
Category	WACV22				CVPR23				CVPR23				CVPR24				AAAI24				NeurIPS24				Ours			
candle	98.8	13.9	23.0	93.8	96.4	10.0	19.9	88.4	77.9	0.5	2.3	55.8	51.8	9.8	5.2	28.5	97.3	12.8	22.8	89.4	**99.0**	**23.1**	**32.4**	95.5	**99.0**	20.7	29.9	**95.7**
capsules	94.1	26.4	30.4	64.3	95.5	42.0	45.5	66.1	86.9	3.3	8.7	62.4	65.9	31.8	34.8	31.8	97.3	10.0	21.0	77.9	**99.1**	**61.3**	**59.8**	**91.8**	99.0	53.9	51.8	91.5
cashew	99.0	53.8	54.9	94.6	98.6	67.8	65.0	82.0	49.7	0.6	2.2	15.7	51.0	22.4	3.1	19.7	90.9	53.1	60.9	61.8	94.3	46.8	51.4	**87.8**	**95.5**	**51.8**	**54.7**	**87.8**
chewinggum	99.1	60.3	60.3	87.9	98.3	31.5	36.9	75.1	68.5	1.3	1.1	39.2	72.4	40.0	42.0	36.4	94.7	11.9	25.8	59.5	98.1	57.5	61.1	79.7	**98.2**	**59.9**	**62.6**	**80.1**
fryum	97.5	53.6	53.3	87.4	94.9	46.2	48.4	81.7	81.7	12.7	19.0	62.7	53.8	35.7	10.4	19.2	97.6	58.6	60.1	81.3	96.9	47.8	51.9	91.6	**97.1**	**48.5**	**52**	**91.7**
macaroni1	99.1	7.6	13.3	94.9	97.0	3.8	10.6	87.0	81.1	0.1	0.2	39.9	60.1	13.2	19.7	27.5	94.1	10.2	16.7	68.5	**99.5**	17.5	**27.6**	95.2	**99.5**	17.1	26.1	**96.1**
macaroni2	97.2	1.5	5.5	88.2	90.9	0.7	4.2	79.1	63.9	0.1	0.1	8.3	51.4	5.3	3.6	16.0	93.6	0.9	2.8	73.1	99.5	9.2	16.1	96.2	**99.6**	**11.4**	**18.7**	**96.8**
pcb1	99.1	72.4	68.7	87.8	98.7	80	74.9	78.9	92.8	49.3	45.5	40.5	70.5	36.9	41.9	25.1	98.7	49.6	52.8	80.2	**99.8**	**77.1**	**72.4**	92.8	**99.8**	76.4	70.7	**93.7**
pcb2	96.6	12.5	18.7	82.3	96.6	13.9	23.5	80.6	93.9	8.8	15.9	71.9	64.1	12.9	19.8	32.7	95.2	7.5	16.7	67.0	**98.9**	**13.3**	**23.4**	**89.6**	98.8	12.4	23.2	89.5
pcb3	96.3	22.6	26.8	80.3	97.6	19.7	27.7	80.8	56.4	0.5	2.7	7.6	71.5	29.8	36.6	35.7	96.7	8.0	18.8	68.9	**99.1**	18.3	27.4	89.1	**99.1**	**21.2**	**27.5**	**90.5**
pcb4	96.7	21.9	30.9	85.4	95.2	22.1	30.9	79.0	89.5	5.0	8.9	66.3	58.3	28.0	23.1	24.0	97.0	17.6	27.2	85.0	98.6	**47**	**46.9**	87.6	**98.7**	45.8	46.6	**89.1**
pipe_fryum	99.2	60.8	61.2	94.6	98.8	70.8	66.2	71.3	81.3	4.7	9.1	43.0	61.3	43.2	31.4	32.3	**99.4**	**72.7**	**69.9**	89.9	99.1	53.5	58.5	**95.1**	99.2	62.3	61.4	93.9
Average	97.7	33.9	37.2	86.8	96.6	34.0	37.8	79.2	77.0	7.2	9.6	42.8	61.0	25.7	22.6	27.4	96.0	26.1	33.0	75.2	98.5	39.4	**44**	91.0	**98.6**	**40.1**	43.8	**91.4**

Table 10. **Comparison of Pixel-level AU-ROC/AP/F1_max/AU-PRO metrics with SoTA on the Visa dataset.** The best results are shown in **bold**.

Method	CFLOW-AD			SimpleNet			PyramidalFlow			RealNet			MambaAD			Wave-MambaAD		
Category	WACV22			CVPR23			CVPR23			CVPR24			NeurIPS24			Ours		
bracket_black	60.8	70.3	75.8	74.0	83.3	78.5	65.8	72.7	77.0	63.0	72.7	75.2	**81.9**	**88.3**	**81.7**	79.7	86.0	81.1
bracket_brown	74.4	77.6	85.2	90.4	93.8	93.6	63.7	76.4	81.0	91.0	94.0	92.6	**95.7**	**97.5**	95.3	92.5	94.6	**95.3**
bracket_white	62.9	64.1	69.8	85.6	89.3	80.0	76.8	84.8	78.4	77.3	86.2	80.8	**97.1**	**97.5**	**90.9**	93.0	93.8	86.7
connector	91.2	85.3	78.6	96.2	92.0	90.3	69.0	42.7	63.4	99.0	97.9	96.6	**99.5**	**99.1**	**96.3**	99.3	98.6	93.3
metal_plate	99.9	100.0	99.3	100.0	100.0	100.0	96.7	98.7	95.2	99.9	100.0	99.3	99.8	99.9	98.6	**100.0**	**100.0**	**100.0**
tubes	65.0	83.3	81.7	84.0	93.9	85.0	69.8	86.6	81.2	80.1	90.1	85.2	58.4	77.0	81.7	**90.5**	**96.3**	**89.6**
Average	75.7	80.1	81.7	88.4	92.0	87.9	73.6	77.0	79.4	85.1	90.2	88.3	88.7	93.2	90.8	**92.5**	**94.9**	**91.0**

Table 11. **Comparison of image-level AU-ROC/AP/F1_max metrics with SoTA on the MPDD dataset.** The best results are shown in **bold**.

Method	CFLOW-AD				SimpleNet				PyramidalFlow				RealNet				MambaAD				Wave-MambaAD			
Category	WACV22				CVPR23				CVPR23				CVPR24				NeurIPS24				Ours			
bracket_black	95.5	1.7	4.1	85.2	92.8	3.1	9.7	86.1	95.3	1.8	5.0	81.4	65.9	2.3	6.8	50.7	94.5	5.3	11.2	88.4	**94.8**	**7.2**	**15.2**	**88.5**
bracket_brown	95.9	5.9	12.0	91.1	94.4	7.9	16.9	86.7	94.0	5.7	12.5	69.6	69.9	12.1	19.1	50.5	**97.5**	**26.3**	**30.8**	92.8	97.4	14.2	23.3	**93.2**
bracket_white	98.2	1.2	3.5	93.0	97.7	2.1	6.5	85.6	98.9	5.6	14.5	91.1	85.2	18.4	30.6	47.4	**99.3**	**12.2**	**20.4**	**95.0**	99.1	2.1	4.9	94.4
connector	96.7	27.9	29.3	89.4	98.3	46.4	47.0	94.5	90.7	3.7	7.2	70.0	85.8	38.9	44.1	75.8	99.2	54.5	56.5	97.3	**99.5**	**69.3**	**61.9**	**98.1**
metal_plate	98.5	89.0	83.3	92.4	98.4	89.8	81.7	89.2	88.7	61.3	21.1	60.8	97.7	90.5	84.1	93.6	98.2	83.8	82.2	**94.3**	**98.8**	**90.6**	**85.5**	85.6
tubes	96.2	32.0	35.6	86.0	97.6	42.8	45.8	91.7	97.1	48.2	46.4	90.3	95.3	54.3	53.0	90.4	96.1	19.6	27.8	85.8	**98.2**	**65.9**	**62.4**	**92.9**
Average	96.8	26.3	28.0	89.5	96.5	32.0	34.6	89.0	94.1	21.1	17.8	77.2	83.3	36.1	39.6	68.1	97.5	33.6	38.2	92.3	**98.0**	**41.5**	**42.2**	**93.8**

Table 12. **Comparison of Pixel-level AU-ROC/AP/F1_max/AU-PRO metrics with SoTA on the MPDD dataset.** The best results are shown in **bold**.

Method	CFLOW-AD			SimpleNet			PyramidalFlow			DiAD			MambaAD			Wave-MambaAD		
Category	WACV22			CVPR23			CVPR23			AAAI24			NeurIPS24			Ours		
bagel	78.8	94.2	89.3	80.7	94.6	89.8	72.4	92.3	88.9	**100.0**	**100.0**	**100.0**	93.4	98.3	93.3	90.4	97.1	93.5
cable_gland	69.7	90.2	90.1	84.3	95.3	91.9	59.3	87.0	89.2	68.1	91.0	92.3	**91.1**	**97.8**	**93.6**	88.6	97.1	92.0
carrot	85.2	96.8	92.3	78.4	95.0	91.2	67.3	88.0	91.9	**94.4**	**99.3**	**98.0**	90.1	97.7	93.9	89.2	97.3	94.1
cookie	46.3	78.0	88.4	62.7	87.2	88.0	15.6	64.7	88.0	**69.4**	78.8	90.9	58.3	85.6	**88.4**	67.3	**89.6**	88.1
dowel	92.1	98.1	92.9	89.7	97.5	90.7	74.5	92.9	89.9	**98.0**	**99.3**	**97.3**	96.5	99.2	95.3	96.0	99.0	94.8
foam	75.0	93.5	89.3	80.0	94.9	89.8	66.5	89.9	88.9	**100.0**	**100.0**	**100.0**	79.9	94.7	89.9	83.5	95.8	89.9
peach	72.2	91.3	90.4	66.5	88.0	90.5	67.2	89.5	89.8	58.0	91.3	**94.3**	89.3	97.1	93.4	**91.3**	**97.6**	75.5
potato	62.0	86.8	89.3	61.6	84.6	89.3	75.0	92.0	90.5	**76.3**	**94.3**	**95.0**	59.1	86.0	90.2	65.1	88.3	70.5
rope	94.1	97.7	92.2	92.9	97.5	93.2	67.5	81.8	81.2	89.2	95.4	91.9	96.1	98.4	93.2	**96.5**	**98.7**	**95.5**
tire	56.1	83.9	87.4	59.7	84.0	88.3	52.5	82.7	87.4	**92.7**	**98.9**	**95.8**	87.3	95.9	91.1	82.9	94.6	89.8
Average	73.1	91.0	90.2	75.7	91.9	90.3	61.8	86.1	88.6	84.6	94.8	**95.6**	84.1	95.1	92.2	**85.1**	**95.5**	92.2

Table 13. **Comparison of image-level AU-ROC/AP/F1_max metrics with SoTA on the MVTec-3D dataset.** The best results are shown in **bold**.

Method	CFLOW-AD				SimpleNet				PyramidalFlow				DiAD				MambaAD				Wave-MambaAD			
Category	WACV22				CVPR23				CVPR23				AAAI24				NeurIPS24				Ours			
bagel	98.3	31.3	38.9	89.7	96.1	21.1	30.6	75.8	95.8	10.4	17.4	84.2	98.5	**49.6**	**54.2**	93.8	98.7	41.0	45.3	94.6	**98.8**	41.8	45.1	**95.4**
cable_gland	96.7	9.4	16.2	90.0	97.9	21.0	27.6	92.7	86.7	1.2	2.7	60.0	98.4	25.2	32	94.5	**99.3**	**38.8**	**44.6**	**98.2**	99.1	37.3	43.0	97.6
carrot	98.9	20.0	27.2	96.0	97.9	12.1	19.7	91.5	98.5	18.0	25.1	94.4	98.6	20	26.9	94.6	99.3	**29.0**	**32.7**	**97.9**	**99.4**	26.9	32.3	97.8
cookie	96.9	30.3	32.6	89.2	93.5	27.6	34.6	74.8	91.5	10.9	18.1	76.0	94.3	14	23.8	83.5	96.8	38.8	42.0	84.3	**97.4**	**40.5**	**43.8**	**87.8**
dowel	98.8	28.3	32.7	94.4	98.2	20.0	24.7	91.7	96.4	15.6	21.5	84.3	97.2	31.4	40.1	89.6	99.5	**50.1**	48.9	**97.4**	**99.6**	49.9	**50.6**	97.1
foam	85.5	15.8	27.2	57.5	87.0	12.2	22.2	69.2	76.7	10.5	17.1	55.7	89.8	9.6	23.5	69.1	**94.9**	**23.8**	**33.4**	**83.8**	94.7	**23.8**	33.2	83.0
peach	97.9	17.1	18.9	92.1	94.0	5.9	11.7	78.5	97.2	8.4	13.0	90.2	98.4	27.6	31.3	94.2	**99.4**	42.9	44.0	97.5	**99.4**	**46.4**	**47.0**	**97.6**
potato	98.5	11.4	13.9	95.3	95.5	2.5	5.7	84.3	98.4	11.1	18.5	94.6	98.0	8.6	17.8	93.9	99.0	17.7	22.8	**95.4**	**99.1**	**18.5**	**24.0**	95.2
rope	99.3	45.1	47.3	95.3	99.5	54.2	53.3	95.2	85.6	7.8	12.3	57.3	99.3	**61.0**	**59.9**	**96.5**	99.3	45.1	48.1	96.1	**99.4**	50.9	50.5	95.5
tire	97.5	7.0	11.2	90.7	93.6	7.9	14.3	76.2	92.9	5.5	11.6	75.6	91.8	5.9	13.7	68.8	99.3	**42.1**	**46.3**	**96.7**	**99.4**	38.2	45.9	96.6
Average	96.8	21.6	26.6	89.0	95.3	18.5	24.5	83.0	92.0	9.9	15.7	77.2	96.4	25.3	32.3	87.8	**98.6**	36.9	40.8	94.2	**98.6**	**37.4**	**41.6**	**94.4**

Table 14. **Comparison of Pixel-level AU-ROC/AP/F1_max/AU-PRO metrics with SoTA on the MVTec-3D dataset.** The best results are shown in **bold**.

Method	CFLOW-AD			SimpleNet			PyramidalFlow			DiAD			MambaAD			Wave-MambaAD		
Category	WACV22			CVPR23			CVPR23			AAAI24			NeurIPS24			Ours		
audiojack	77.4	70.8	60.9	58.4	44.2	50.9	51.7	34.9	50.7	76.5	66.0	65.7	84.9	77.7	68.2	**85.5**	**80.4**	**68.6**
bottle cap	80.0	78.3	71.7	58.2	47.6	45.2	55.4	46.7	60.3	91.6	87.0	87.9	93.2	92.6	82.7	**95.4**	**95.3**	**88.3**
button battery	66.0	75.1	72.7	77.2	60.5	77.6	52.5	56.4	72.5	80.5	54.3	70.6	82.8	87.4	79.2	**88.0**	**90.0**	**84.0**
end cap	63.7	72.7	72.9	54.1	60.8	60.3	55.7	63.8	73.0	85.1	94.0	84.8	78.6	83.1	77.4	**81.5**	**85.5**	**78.4**
eraser	88.6	87.5	77.6	52.5	39.1	72.4	57.8	42.5	57.3	80.0	71.3	77.3	88.4	86.9	76.7	**91.2**	**89.3**	**79.1**
fire hood	80.3	73.8	68.2	51.6	41.9	72.9	56.5	39.1	54.1	83.3	83.4	80.5	79.9	73.5	65.4	**83.5**	**77.7**	**69.5**
mint	63.6	64.1	63.7	46.4	50.3	55.8	57.0	51.2	63.8	76.7	80.0	76.0	72.6	73.8	66.2	**76.4**	**77.4**	**68.1**
mounts	82.9	74.5	70.3	58.1	48.1	54.4	56.1	41.1	52.1	75.3	81.7	82.5	87.4	78.3	74.2	**88.1**	**80.6**	**75.2**
pcb	74.3	83.2	76.6	52.4	66.0	63.7	54.8	64.5	75.7	86.0	76.7	85.4	90.3	94.3	85.1	**90.9**	**94.5**	**86.0**
phone battery	74.9	73.4	64.2	58.7	43.8	52.4	45.5	36.1	58.4	82.3	74.5	75.9	90.2	88.6	80.6	**91.4**	**89.6**	**81.4**
plastic nut	69.7	60.5	53.6	54.5	40.3	75.5	45.0	29.3	49.7	71.9	85.1	65.6	87.8	82.0	71.9	**89.0**	**82.8**	**73.9**
plastic plug	78.8	75.2	64.9	51.6	38.4	58.0	46.2	35.2	54.6	88.7	77.7	90.9	86.5	83.4	72.8	**87.3**	**83.8**	**73.8**
procelain doll	83.6	78.1	68.2	59.2	54.5	51.8	48.7	34.0	49.8	72.6	58.2	65.2	**88.4**	**82.7**	**74.4**	86.0	78.8	70.4
regulator	50.5	29.5	43.9	48.2	29.0	54.6	55.1	31.3	44.9	72.1	89.2	78.2	72.1	62.7	53.4	**80.9**	**71.9**	**61.9**
rolled strip base	92.6	96.6	88.9	66.3	75.7	52.1	59.9	74.9	79.8	68.4	66.8	56.8	98.3	99.2	95.6	**98.8**	**99.4**	**96.4**
sim card set	91.5	92.9	85.3	50.5	69.7	43.9	77.9	75.3	77.1	72.6	71.4	61.5	**94.7**	95.4	**87.9**	94.6	**95.5**	87.5
switch	75.3	79.9	72.5	59.0	66.8	79.8	60.8	62.0	69.9	73.4	55.9	61.2	92.4	94.3	**86.3**	**93.1**	**95.0**	**86.3**
tape	93.5	92.6	84.3	63.1	41.1	70.8	60.4	43.3	58.3	73.9	53.7	66.1	**97.1**	**96.2**	**89.6**	**97.1**	**96.2**	89.4
terminalblock	81.1	84.3	76.0	62.2	64.7	68.6	57.8	57.5	70.0	62.1	49.4	47.8	95.3	95.5	89.8	**96.7**	**97.2**	**91.2**
toothbrush	70.3	74.9	71.7	49.9	70.0	54.5	48.1	50.9	70.1	91.2	57.8	90.9	**86.2**	**87.5**	80.7	85.6	86.2	**81.4**
toy	60.6	68.8	73.7	59.8	64.4	68.8	56.2	65.2	73.4	66.2	36.4	59.8	83.7	88.4	79.7	**85.8**	**89.1**	**82.6**
toy brick	74.2	69.7	64.0	65.9	49.7	70.1	54.6	43.4	58.2	68.4	93.7	55.9	70.6	64.0	61.8	**77.7**	**72.7**	**66.9**
transisitor1	91.8	94.7	86.2	57.8	69.2	73.4	49.1	56.2	72.4	73.1	57.3	62.7	94.9	96.4	89.2	**96.8**	**97.6**	**91.9**
u block	80.3	73.0	63.8	58.3	48.4	58.2	44.9	28.6	48.8	75.2	45.3	67.9	90.0	85.8	74.8	**91.6**	**88.4**	**79.8**
usb	68.0	69.8	63.4	62.2	55.3	72.1	48.3	44.5	63.0	58.9	63.1	45.7	**92.7**	**92.7**	**85.3**	92.6	92.2	84.7
usb adaptor	69.9	64.1	59.5	62.4	38.4	51.8	53.4	40.5	56.8	76.9	68.4	67.2	79.1	75.7	66.0	**81.8**	**76.5**	**68.8**
vcpill	81.4	79.1	67.5	57.0	48.7	62.9	55.7	43.9	57.8	64.1	37.4	56.2	88.5	87.6	77.6	**90.7**	**90.0**	**80.6**
woodstick	79.7	79.3	68.6	47.5	52.0	56.5	66.4	56.3	64.8	62.1	60.2	65.9	83.3	82.3	72.7	**86.3**	**85.4**	**75.8**
wooden beads	71.3	60.5	52.8	59.0	35.6	56.4	50.1	28.1	44.4	74.1	40.4	62.1	81.8	71.0	64.4	**84.4**	**77.3**	**68.0**
zipper	94.3	97.1	90.3	55.1	86.7	60.2	49.2	61.7	77.5	86.0	56.4	84.0	**99.4**	**99.6**	**97.1**	99.3	**99.6**	96.8
Average	77.0	75.8	69.9	57.2	53.4	61.5	54.5	47.9	62.0	75.6	66.4	69.9	87.0	85.3	77.6	**88.9**	**87.2**	**79.6**

Table 15. **Comparison of image-level AU-ROC/AP/F1_max metrics with SoTA on the Real-IAD dataset.** The best results are shown in **bold**.

Method	CFLOW-AD				SimpleNet				PyramidalFlow				DiAD				MambaAD				Wave-MambaAD			
Category	WACV22				CVPR23				CVPR23				AAAI24				NeurIPS24				Ours			
audiojack	95.4	16.4	15.4	73.2	66.0	0.2	0.9	37.6	81.7	0.2	0.4	40.7	91.6	1.0	3.9	63.3	98.0	22.2	30.4	86.2	**98.1**	**35.1**	**44.4**	**87.7**
bottle cap	98.8	13.6	22.6	91.9	80.1	0.5	3.7	35.8	93.2	0.7	2.6	72.0	94.6	4.9	11.4	73.0	**99.7**	29.9	33.8	97.4	99.6	**38.2**	**38.7**	96.8
button battery	95.6	29.2	25.8	74.8	78.0	9.4	15.4	38.3	50.2	17.5	0.8	15.0	84.1	1.4	5.3	66.9	**98.3**	48.6	49.6	**88.1**	97.8	**51.6**	**52.9**	85.8
end cap	87.6	3.5	5.9	60.3	65.5	0.3	1.8	39.0	77.2	0.4	1.9	33.2	81.3	2.0	6.9	38.2	97.2	12.2	19.5	90.3	**99**	**35.6**	**40.6**	**93.3**
eraser	98.9	23.2	26.9	93.4	87.6	3.8	8.7	59.8	88.8	0.7	2.6	60.3	91.1	7.7	15.4	67.5	**99.2**	27.4	35.7	**93.8**	98.4	**34.7**	**39.7**	91.1
fire hood	98.1	21.6	22.7	86.0	76.8	1.2	4.8	39.9	50.0	0.2	0.1	15.4	91.8	3.2	9.2	66.7	**98.7**	25.9	33.0	86.8	98.6	**30.0**	**36.4**	**89.9**
mint	92.8	8.7	12.8	61.4	73.1	0.5	3.0	26.8	50.0	0.1	0.1	15.0	91.1	5.7	11.6	64.2	**97.1**	16.7	27.2	76.9	97	**22.7**	**32.7**	**79.8**
mounts	96.1	16.9	25.5	86.4	87.2	1.4	4.2	63.0	82.7	1.0	4.0	54.3	84.3	0.4	1.1	48.8	**99.1**	30.9	35.2	**92.6**	**99.1**	**37.0**	**39.3**	91.8
pcb	95.1	15.0	23.8	75.6	76.8	0.4	1.0	45.8	81.6	0.3	0.7	45.0	92.0	3.7	7.4	66.5	**99.2**	48.4	51.7	**93.8**	99.1	**51.1**	**53.8**	92.7
phone battery	75.2	13.6	22.7	83.9	75.4	1.5	6.0	46.8	71.6	0.3	1.2	27.2	96.8	5.3	11.4	85.4	**99.4**	**35.2**	**39.9**	**95.5**	98.9	33.8	39.0	94.7
plastic nut	95.6	13.4	14.9	79.1	74.0	0.5	2.3	40.9	76.3	0.1	0.4	29.3	81.1	0.4	3.4	38.6	**99.5**	**34.3**	**38.0**	**96.8**	99.4	34.2	37.9	96.3
plastic plug	97.2	11.8	20.7	85.7	75.9	0.2	0.8	38.6	79.7	0.1	0.3	40.5	92.9	8.7	15.0	66.1	**99.0**	**25.2**	**32.3**	**93.0**	98.4	23.1	31.4	89.1
procelain doll	97.1	14.0	20.9	86.6	81.4	2.4	7.4	48.9	84.5	0.2	0.3	48.7	93.1	1.4	4.8	70.4	**99.2**	**32.2**	**36.9**	**95.9**	98.4	29.3	34.8	91.2
regulator	88.2	1.2	2.1	58.0	79.9	0.1	0.7	40.8	76.9	0.1	0.3	42.9	84.2	0.4	1.5	44.4	97.9	21.7	30.1	88.5	**98.3**	**26.5**	**33.3**	**90**
rolled strip base	97.8	10.6	15.0	93.2	77.7	1.5	4.9	50.3	88.1	1.2	3.5	69.1	87.7	0.6	3.2	63.4	99.6	27.4	32.3	98.2	**99.7**	**36.9**	**43.2**	**98.9**
sim card set	98.2	30.3	35.2	87.9	73.9	2.8	7.1	30.3	66.0	1.0	0.6	31.0	89.9	1.7	5.8	60.4	**98.7**	51.0	50.4	**89.1**	98.1	**51.4**	**50.8**	88
switch	90.9	13.9	17.7	78.8	69.9	1.2	3.1	49.3	50.0	0.3	0.6	15.0	90.5	1.4	5.3	64.2	98.4	34.9	42.2	93.8	**99.2**	**55.2**	**56.5**	**94.4**
tape	99.2	24.6	24.8	95.6	82.8	1.3	4.0	42.8	66.9	0.2	0.4	10.6	81.7	0.4	2.7	47.3	**99.8**	**45.9**	**48.4**	**98.4**	99.7	45	47.6	98.2
terminalblock	97.0	12.1	17.7	86.5	85.0	0.7	2.0	57.7	91.0	0.5	1.7	67.3	75.5	0.1	1.1	38.5	99.6	26.8	32.8	97.6	**99.8**	**40.7**	**44.2**	**98.2**
toothbrush	94.7	18.8	14.0	78.0	79.1	3.0	7.3	46.3	50.7	8.4	2.6	15.6	82.0	1.9	6.6	54.5	**97.6**	30.1	37.9	**91.9**	97.5	**32.2**	**39.1**	90.9
toy	87.5	2.1	8.2	56.9	75.2	0.2	0.9	36.9	79.8	0.4	1.6	47.4	82.1	1.1	4.2	50.3	96.2	**16.8**	**26.1**	88.0	**96.5**	16.3	25.3	**88.7**
toy brick	96.1	24.2	26.7	79.2	85.9	4.5	10.6	47.0	50.0	0.1	0.3	35.3	93.5	3.1	8.1	66.4	96.6	18.9	26.8	75.9	**97.7**	**27.2**	**34.5**	**86.4**
transisitor1	98.1	26.2	28.3	90.7	84.5	5.4	10.1	58.5	77.0	0.4	0.8	39.5	88.6	7.2	15.3	58.1	**99.4**	38.1	39.4	96.8	**99.4**	**43.1**	**42.6**	**97**
u block	98.4	19.9	24.7	89.3	72.8	0.6	2.5	40.2	81.9	0.5	1.1	39.8	88.8	1.6	5.4	54.2	**99.5**	33.2	42.8	**96.1**	99.4	**40.7**	**47.7**	95.2
usb	94.8	13.0	15.8	75.0	80.4	0.8	2.8	50.8	79.2	0.2	0.2	40.0	78.0	1.0	3.1	28.0	99.3	39.7	44.8	96.0	**99.4**	**42.8**	**48.1**	**96.2**
usb adaptor	96.2	8.1	16.6	80.8	52.3	0.1	0.4	19.2	84.2	0.2	0.6	48.4	94.0	2.3	6.6	75.5	97.0	15.8	24.9	81.6	**97.1**	**21.5**	**30.5**	**82.3**
vcpill	97.3	35.6	41.3	84.6	82.0	6.5	12.0	47.6	72.3	0.6	0.8	28.4	90.2	1.3	5.2	60.8	98.7	47.9	52.0	89.4	**98.8**	**52.8**	**55.4**	**91.3**
woodstick	96.5	20.2	27.3	79.5	75.5	1.2	4.4	36.2	50.1	0.6	0.3	15.1	85.0	1.1	4.7	45.6	**98.1**	32.4	39.7	85.3	98.0	**44.2**	**49.6**	**89**
wooden beads	92.6	32.1	38.3	67.9	74.0	2.8	9.1	32.1	50.0	0.1	0.2	15.0	90.9	2.6	8.0	60.7	97.9	42.8	47.0	85.5	**98.0**	**44.2**	**49.6**	**89**
zipper	98.2	34.9	35.6	91.0	52.7	1.2	3.8	24.3	50.0	0.7	1.4	15.6	90.2	12.5	18.8	53.5	**99.3**	**58.4**	**60.9**	**97.6**	99.0	54.5	55.6	96.5
Average	94.8	17.6	21.7	80.4	76.1	1.9	4.9	42.4	71.1	1.2	1.1	35.8	88.0	2.9	7.1	58.0	**98.6**	32.4	38.1	91.2	**98.6**	**37.7**	**42.5**	**91.7**

Table 16. **Comparison of Pixel-level AU-ROC/AP/F1_max/AU-PRO metrics with SoTA on the MVTec-3D dataset.** The best results are shown in **bold**.