

Table 1: Anomaly detection (left) and location (right) results with AUROC on MVTec-AD. All methods are evaluated under the one-for-all / one-for-one settings. Under the one-for-all setting, the learned model is applied to detect anomalies for all categories without fine-tuning. For fairness, we replace the backbone of PatchCore and CSFLOW with EfficientNet, in accordance with UniAD and our model. We achieves SOTA or competitive results compared with the baselines.

Category		SimpleNet (CVPR'23)		RD++ (CVPR'23)		PatchCore (CVPR'22)		RD4AD (CVPR'22)		CSFLOW (WCCV'22)		UTRAD (NN'22)		Ours	
Object	Bottle	97.7 / 100	91.2 / 98.0	99.5 / 100	98.4 / 98.8	100 / 100	97.4 / 97.9	98.7 / 100	97.7 / 98.7	- / 99.8	- / -	100 / 100	96.4 / 95.9	100 ± 0.0 / 99.9	98.3 ± 0.04 / 98.2
	Cable	87.6 / 99.9	88.1 / 97.6	93.8 / 99.2	90.2 / 98.4	95.3 / 97.8	93.6 / 96.9	85.0 / 95.0	83.1 / 97.4	- / 99.1	- / -	97.8 / 98.6	97.1 / 97.3	99.0 ± 0.29 / 97.8	98.1 ± 0.04 / 97.4
	Capsule	78.3 / 97.7	89.7 / 98.9	86.2 / 99.0	98.4 / 98.8	96.8 / 96.7	98.0 / 98.8	95.5 / 96.3	98.5 / 98.7	- / 97.1	- / -	82.0 / 94.3	97.2 / 97.8	95.4 ± 0.21 / 88.3	98.8 ± 0.01 / 98.6
	Hazelnut	99.2 / 100	95.7 / 97.9	100 / 100	99.1 / 99.2	99.3 / 99.7	97.6 / 97.7	87.1 / 99.9	98.7 / 98.9	- / 99.6	- / -	99.8 / 99.5	98.2 / 98.4	100 ± 0.08 / 100	98.8 ± 0.02 / 98.5
	Metal Nut	85.1 / 100	90.9 / 98.8	99.9 / 100	97.1 / 98.1	99.1 / 99.8	96.3 / 96.7	99.4 / 100	94.1 / 97.3	- / 99.1	- / -	94.7 / 96.2	96.4 / 95.0	99.9 ± 0.02 / 99.4	96.3 ± 0.09 / 95.0
	Pill	78.3 / 99.0	89.7 / 98.6	96.7 / 98.4	98.4 / 98.3	86.4 / 85.5	90.8 / 91.9	52.6 / 96.6	96.5 / 98.2	- / 98.6	- / -	89.7 / 94.2	95.7 / 97.5	95.8 ± 0.49 / 94.2	97.1 ± 0.04 / 95.6
	Screw	45.5 / 98.2	93.7 / 99.3	95.0 / 98.9	99.4 / 99.7	94.2 / 94.8	98.9 / 99.2	97.3 / 97.0	99.4 / 99.6	- / 97.6	- / -	75.1 / 88.3	95.2 / 97.8	95.6 ± 0.60 / 93.1	98.9 ± 0.05 / 98.9
	Toothbrush	94.7 / 99.7	97.5 / 98.5	99.2 / 100	99.1 / 99.1	100 / 100	98.8 / 98.8	99.4 / 99.5	99.0 / 99.1	- / 91.9	- / -	89.7 / 78.9	97.5 / 96.2	93.6 ± 0.66 / 92.2	98.6 ± 0.04 / 98.4
	Transistor	82.0 / 100	86.0 / 97.6	94.2 / 98.5	89.6 / 94.3	98.9 / 99.6	92.3 / 94.1	92.4 / 96.7	86.4 / 92.5	- / 99.3	- / -	92.0 / 96.4	91.5 / 94.9	99.7 ± 0.12 / 99.1	97.9 ± 0.04 / 98.1
	Zipper	99.1 / 99.9	97.0 / 98.9	99.5 / 98.6	98.1 / 98.8	97.1 / 98.9	95.7 / 97.5	99.6 / 98.5	98.1 / 98.2	- / 99.7	- / -	95.5 / 98.6	97.3 / 97.9	97.9 ± 0.15 / 92.5	97.5 ± 0.10 / 98.0
Texture	Carpet	95.9 / 99.7	92.4 / 98.2	96.5 / 100	98.9 / 99.2	97.0 / 95.9	98.1 / 98.0	97.1 / 98.9	98.8 / 98.9	- / 100	- / -	80.3 / 96.3	94.4 / 97.3	99.9 ± 0.03 / 100	98.7 ± 0.03 / 98.7
	Grid	49.8 / 99.7	46.7 / 98.8	99.2 / 100	98.9 / 99.3	91.4 / 92.5	98.4 / 98.3	99.7 / 100	99.2 / 99.3	- / 99.0	- / -	93.9 / 98.7	95.2 / 97.6	97.0 ± 0.69 / 98.7	97.0 ± 0.06 / 95.5
	Leather	93.9 / 100	96.9 / 99.2	99.9 / 100	99.4 / 99.4	100 / 100	99.2 / 99.2	100 / 100	99.4 / 99.4	- / 100	- / -	99.8 / 100	98.4 / 98.6	100 ± 0.00 / 100	98.8 ± 0.02 / 99.1
	Tile	93.7 / 99.8	93.1 / 97.0	99.5 / 99.7	97.0 / 96.6	96.0 / 95.9	90.3 / 90.8	97.5 / 99.3	95.6 / 95.6	- / 100	- / -	98.8 / 99.9	94.2 / 95.0	99.2 ± 0.32 / 99.2	92.2 ± 0.42 / 92.0
	Wood	95.2 / 100	84.8 / 94.5	98.5 / 99.3	95.8 / 95.8	93.8 / 95.1	90.8 / 91.0	99.2 / 99.2	96.0 / 95.9	- / 100	- / -	99.7 / 99.7	89.4 / 93.1	97.2 ± 0.40 / 98.6	92.4 ± 0.16 / 93.9
Mean		85.1 / 99.6	88.9 / 98.1	97.2 / 99.4	97.2 / 98.3	96.4 / 96.8	95.7 / 96.5	93.4 / 98.5	96.0 / 97.8	- / 98.7	- / -	92.6 / 96.0	95.6 / 96.7	98.0 ± 0.11 / 96.9	97.3 ± 0.05 / 97.1

Table 2: Complexity analysis including inference FLOPs and learnable parameters with the input image size of 224×224 . The results show that the advantage of our approach does not come from a larger model capacity, and the complexity of our model is competitive when achieving superior performance.

	US (CVPR'20)	PSVDD (ACCV'20)	PaDiM (ICPR'21)	CutPaste (CVPR'21)	FCDD (ICLR'21)	MKD (CVPR'21)	DREAM (ICCV'21)	UniAD (NeurIPS'21)	Ours
Params(M)	9.55	0.41	950.36	13.61	4.51	0.34	69.05	7.48	7.54
FLOPs(G)	60.32	149.74	23.25	3.65	13.16	32.11	245.15	6.46	6.60

Table 3: Anomaly detection (left) and location (right) results with AUROC metric on VisA, a more complicated dataset than MVTec-AD, evaluated under the one-for-all / one-for-one settings. Our model is applied to all categories without specific parameter-tuning on each category.

Category		DRAEM (CVPR'23)		UniAD (CVPR'23)		Ours	
Complex structure	PCB1	83.9 / 71.3	94.0 / 98.6	95.4 / 89.7	99.3 / 99.2	96.7 / 93.5	99.4 / 99.0
	PCB2	81.7 / 89.7	94.1 / 92.5	93.6 / 85.3	97.8 / 96.5	93.4 / 92.0	98.0 / 97.5
	PCB3	87.7 / 73.1	94.1 / 93.8	88.6 / 89.7	98.3 / 97.9	92.0 / 91.9	98.3 / 98.1
	PCB4	87.1 / 91.3	72.3 / 95.8	99.4 / 97.6	97.9 / 97.0	99.5 / 98.5	97.7 / 96.7
Multiple instances	Macaroni 1	68.6 / 70.3	89.8 / 95.8	92.2 / 86.3	99.3 / 98.6	93.1 / 91.7	99.4 / 99.2
	Macaroni 2	60.3 / 71.3	83.2 / 94.1	85.9 / 72.4	98.0 / 96.1	86.2 / 76.4	98.5 / 97.1
	Capsules	89.6 / 77.3	96.6 / 93.7	72.0 / 78.7	98.3 / 98.2	77.1 / 76.0	99.0 / 98.5
	Candles	70.2 / 82.3	82.6 / 87.0	96.8 / 91.8	99.2 / 98.7	96.8 / 95.7	99.2 / 99.0
Single instance	Cashew	67.3 / 94.2	68.5 / 94.7	92.4 / 93.7	98.7 / 99.2	94.9 / 93.0	99.2 / 99.2
	Chewing gum	90.0 / 93.4	92.7 / 97.5	99.4 / 97.3	99.2 / 98.5	99.4 / 96.9	98.8 / 98.2
	Fryum	86.2 / 100	83.2 / 97.5	89.8 / 81.6	97.7 / 97.1	90.4 / 88.4	97.7 / 97.4
	Pipe Fryum	87.1 / 94.1	72.3 / 81.8	97.4 / 92.2	99.2 / 99.2	98.5 / 90.3	99.4 / 99.2
Mean		80.5 / 84.1	87.0 / 88.8	91.9 / 88.0	98.6 / 98.0	93.2 / 90.4	98.7 / 98.3

Table 4: Extra ablation studies with AUROC metric on MVTec-AD. "w/o VQ" means without VQ.

w/o VQ	with VQ		Switching		POT	Detection	Localization
	Single	Hierarchical	Codebook-Switching	Expert-Switching			
-	-	-	-	-	-	97.1	96.9
✓	-	✓	-	-	✓	97.4	97.2

Table 5: Anomaly detection (left) and location (right) results with AUROC metric on MVTec-AD. Our model is evaluated under the one-for-all / one-for-one settings. Under the one-for-all setting, we randomly choose 5 categories from all 15 categories and repeat it for 3 times, the learned model is applied to detect anomalies for all categories without fine-tuning. The mean AUCs of detection and location are 97.6% and 97.3%, respectively.

Category	Ours		Category	Ours		Category	Ours	
Capsule	93.8 / 71.3	98.8 / 98.6	Zipper	97.5 / 89.7	97.7 / 99.2	Metal nut	99.9 / 93.5	96.7 / 99.0
Wood	97.0 / 89.7	92.3 / 92.5	Hazelnut	99.4 / 85.3	98.6 / 96.5	Screw	92.7 / 92.0	98.6 / 97.5
Transistor	99.6 / 73.1	98.3 / 93.8	Cable	98.0 / 89.7	97.8 / 97.9	Bottle	100 / 91.9	98.4 / 98.1
Pill	95.8 / 91.3	95.8 / 95.8	Toothbrush	93.9 / 98.5	98.6 / 97.0	Leather	100 / 98.6	99.0 / 96.7
Grid	97.7 / 91.3	97.4 / 95.8	Carpet	99.8 / 98.5	98.8 / 97.0	Tile	99.5 / 98.6	92.5 / 96.7
Mean	96.8 / 84.1	96.5 / 88.8	Mean	97.7 / 88.0	98.3 / 98.0	Mean	98.4 / 90.4	97.0 / 98.3

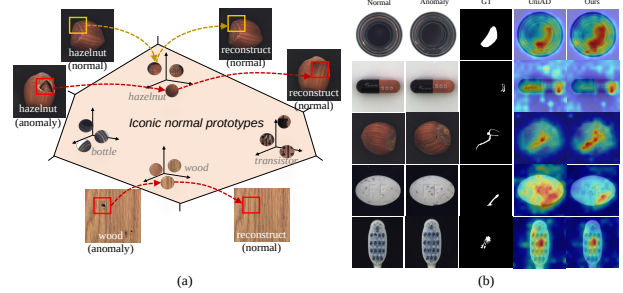


Figure 1: (a) Our modified figure from Fig. 1 of the original submission for showing our motivation; (b) More visualization comparisons between UniAD and our model on MVTec-AD.

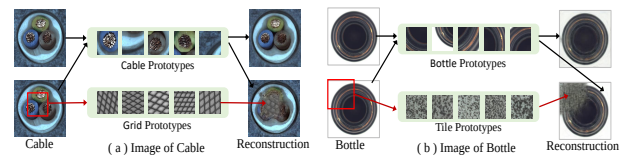


Figure 2: Our modification of Fig. 5 in the original submission. Reconstructed cable results in (a) and reconstructed bottle results in (b) using the correct prototypes or wrong prototypes.