1 Supplemental Materials

1.1 Detailed Comparison on Datasets

1.1.1 MVTec Dataset

Table 1: Detailed performance comparison of NKD on the MVTec dataset.

Category	**NKD**		**NKD+FMAM**	
	Image AUC	Pixel AUC	Image AUC	Pixel AUC
bottle	$100.0_{\pm 0.0}$	$98.7_{\pm 0.0}$	$100.0_{\pm 0.0}$	$98.7_{\pm 0.0}$
cable	$96.1_{\pm 1.0}$	$96.3_{\pm 0.2}$	$95.1_{\pm 1.4}$	$96.1_{\pm 0.2}$
capsule	$91.2_{\pm 1.8}$	$98.4_{\pm 0.1}$	$90.6_{\pm 2.3}$	$98.3_{\pm 0.1}$
carpet	$85.6_{\pm 6.5}$	$97.4_{\pm 1.6}$	$95.9_{\pm 1.1}$	$98.4_{\pm 0.2}$
grid	$99.4_{\pm 0.8}$	$98.7_{\pm 0.2}$	$99.7_{\pm 0.2}$	$98.8_{\pm 0.1}$
hazelnut	$96.5_{\pm 2.3}$	$98.1_{\pm 0.4}$	$99.2_{\pm 0.7}$	$98.0_{\pm 0.2}$
leather	$92.0_{\pm 5.0}$	$97.3_{\pm 0.7}$	$100.0_{\pm 0.0}$	$99.4_{\pm 0.0}$
metal	$99.1_{\pm 0.1}$	$97.1_{\pm 0.1}$	$99.3_{\pm 0.1}$	$97.0_{\pm 0.1}$
pill	$97.2_{\pm 0.3}$	$98.5_{\pm 0.1}$	$96.9_{\pm 0.6}$	$98.4_{\pm 0.1}$
screw	$89.9_{\pm 1.0}$	$98.8_{\pm 0.1}$	$90.6_{\pm 2.3}$	$98.3_{\pm 0.8}$
tile	$91.2_{\pm 4.6}$	$93.6_{\pm 1.7}$	$97.6_{\pm 1.0}$	$94.8_{\pm 0.6}$
toothbrush	$93.3_{\pm 0.9}$	$98.6_{\pm 0.2}$	$93.1_{\pm 0.5}$	$98.6_{\pm 0.3}$
transistor	$95.1_{\pm 0.4}$	$84.6_{\pm 0.1}$	$94.9_{\pm 0.5}$	$84.3_{\pm 0.2}$
wood	$99.3_{\pm 1.0}$	$94.1_{\pm 0.9}$	$99.2_{\pm 0.3}$	$95.6_{\pm 0.6}$
zipper	$95.4_{\pm 1.0}$	$98.1_{\pm 0.2}$	$97.8_{\pm 0.6}$	$98.5_{\pm 0.2}$
Average	$94.7_{\pm 0.6}$	$96.6_{\pm 0.2}$	$96.7_{\pm 0.3}$	$96.9_{\pm 0.1}$

Table 2: Detailed performance comparison of RKD on the MVTec dataset.

Category	**RKD**		**RKD+FMAM**	
	Image AUC	Pixel AUC	Image AUC	Pixel AUC
bottle	$99.9_{\pm 0.1}$	$98.3_{\pm 0.0}$	$100.0_{\pm 0.0}$	$98.3_{\pm 0.1}$
cable	$97.8_{\pm 0.6}$	$96.9_{\pm 0.1}$	$98.5_{\pm 0.6}$	$96.7_{\pm 0.2}$
capsule	$93.4_{\pm 4.6}$	$98.8_{\pm 0.1}$	$94.4_{\pm 1.9}$	$98.9_{\pm 0.1}$
carpet	$99.4_{\pm 0.4}$	$99.0_{\pm 0.0}$	$99.7_{\pm 0.2}$	$98.9_{\pm 0.1}$
grid	$90.0_{\pm 3.6}$	$98.9_{\pm 0.2}$	$99.3_{\pm 0.5}$	$99.0_{\pm 0.1}$
hazelnut	$95.9_{\pm 1.4}$	$98.7_{\pm 0.1}$	$99.9_{\pm 0.2}$	$98.6_{\pm 0.1}$
leather	$100.00_{\pm 0.0}$	$99.3_{\pm 0.0}$	$100.0_{\pm 0.0}$	$99.6_{\pm 0.0}$
metal	$100.00_{\pm 0.0}$	$96.6_{\pm 0.1}$	$100.0_{\pm 0.1}$	$96.5_{\pm 0.2}$
pill	$94.7_{\pm 2.3}$	$97.5_{\pm 0.1}$	$97.8_{\pm 0.4}$	$97.6_{\pm 0.3}$
screw	$84.9_{\pm 4.3}$	$99.1_{\pm 0.1}$	$93.8_{\pm 1.8}$	$98.7_{\pm 0.2}$
tile	$98.4_{\pm 0.1}$	$94.3_{\pm 0.1}$	$97.9_{\pm 0.4}$	$94.5_{\pm 0.2}$
toothbrush	$96.1_{\pm 0.6}$	$99.0_{\pm 0.1}$	$97.6_{\pm 1.0}$	$99.0_{\pm 0.1}$
transistor	$96.2_{\pm 0.3}$	$88.6_{\pm 0.2}$	$96.3_{\pm 0.2}$	$88.4_{\pm 0.2}$
wood	$98.8_{\pm 0.4}$	$94.1_{\pm 0.2}$	$98.5_{\pm 0.2}$	$95.5_{\pm 0.1}$
zipper	$96.1_{\pm 0.4}$	$98.5_{\pm 0.1}$	$99.1_{\pm 0.1}$	$99.0_{\pm 0.1}$
Average	$96.1_{\pm 0.5}$	$97.1_{\pm 0.1}$	$98.2_{\pm 0.1}$	$97.3_{\pm 0.1}$

1.1.2 ZJU-Leaper Dataset

Table 3: Detailed performance comparison of NKD on the ZJU-Leaper dataset.

Category	**NKD**		**NKD+FMAM**	
	Image AUC	Pixel AUC	Image AUC	Pixel AUC
p1	$82.2_{\pm 4.8}$	$91.8_{\pm 2.8}$	$87.2_{\pm 2.2}$	$93.7_{\pm 1.2}$
p2	$98.8_{\pm 0.2}$	$97.9_{\pm 0.1}$	$99.1_{\pm 0.1}$	$98.4_{\pm 0.2}$
p3	$93.0_{\pm 2.5}$	$98.2_{\pm 0.5}$	$98.5_{\pm 0.7}$	$99.1_{\pm 0.4}$
p4	$98.8_{\pm 0.3}$	$96.1_{\pm 0.3}$	$99.3_{\pm 0.3}$	$97.1_{\pm 0.2}$
p5	$94.1_{\pm 2.1}$	$95.8_{\pm 0.8}$	$95.8_{\pm 1.2}$	$96.7_{\pm 0.7}$
p6	$97.0_{\pm 0.3}$	$96.2_{\pm 0.4}$	$95.2_{\pm 1.5}$	$97.1_{\pm 0.4}$
p7	$91.2_{\pm 2.0}$	$91.5_{\pm 1.0}$	$94.2_{\pm 2.4}$	$92.6_{\pm 0.9}$
p8	$86.6_{\pm 4.4}$	$96.0_{\pm 0.9}$	$90.4_{\pm 2.1}$	$97.6_{\pm 0.3}$
p9	$94.3_{\pm 0.6}$	$97.3_{\pm 0.2}$	$95.9_{\pm 1.7}$	$97.9_{\pm 0.5}$
p10	$61.9_{\pm 8.6}$	$84.7_{\pm 5.8}$	$64.2_{\pm 8.8}$	$83.6_{\pm 7.7}$
p11	$90.1_{\pm 4.4}$	$92.6_{\pm 1.7}$	$94.6_{\pm 1.8}$	$95.3_{\pm 0.5}$
p12	$71.0_{\pm 9.0}$	$87.9_{\pm 4.0}$	$66.9_{\pm 15.9}$	$81.3_{\pm 12.1}$
p13	$75.4_{\pm 4.9}$	$90.4_{\pm 0.7}$	$89.3_{\pm 12.6}$	$93.6_{\pm 2.1}$
p14	$72.0_{\pm 7.1}$	$88.5_{\pm 0.5}$	$80.7_{\pm 3.1}$	$91.1_{\pm 1.0}$
p15	$66.5_{\pm 3.6}$	$84.9_{\pm 1.5}$	$77.8_{\pm 5.1}$	$89.5_{\pm 1.3}$
Average	$84.9_{\pm 1.2}$	$92.7_{\pm 0.4}$	$88.6_{\pm 2.3}$	$93.6_{\pm 1.2}$

Table 4: Detailed performance comparison of RKD on the ZJU-Leaper dataset.

Category	**RKD**		**RKD+FMAM**	
	Image AUC	Pixel AUC	Image AUC	Pixel AUC
p1	$97.4_{\pm 0.6}$	$98.9_{\pm 0.1}$	$97.5_{\pm 0.5}$	$98.2_{\pm 0.4}$
p2	$98.2_{\pm 0.3}$	$97.2_{\pm 0.5}$	$98.5_{\pm 0.2}$	$97.8_{\pm 0.2}$
p3	$98.1_{\pm 1.2}$	$99.3_{\pm 0.2}$	$98.8_{\pm 0.4}$	$99.3_{\pm 0.1}$
p4	$98.1_{\pm 0.3}$	$96.6_{\pm 0.0}$	$98.0_{\pm 0.3}$	$96.5_{\pm 0.2}$
p5	$96.6_{\pm 0.2}$	$96.1_{\pm 0.6}$	$96.8_{\pm 0.2}$	$95.7_{\pm 1.0}$
p6	$94.0_{\pm 1.6}$	$97.0_{\pm 0.3}$	$92.9_{\pm 1.3}$	$96.7_{\pm 0.2}$
p7	$87.1_{\pm 0.8}$	$88.0_{\pm 0.2}$	$86.4_{\pm 4.3}$	$89.1_{\pm 0.7}$
p8	$86.3_{\pm 3.6}$	$95.8_{\pm 0.4}$	$93.0_{\pm 2.0}$	$97.6_{\pm 0.3}$
p9	$91.5_{\pm 2.6}$	$97.0_{\pm 0.2}$	$96.0_{\pm 1.1}$	$97.2_{\pm 0.1}$
p10	$81.6_{\pm 1.2}$	$93.2_{\pm 0.3}$	$93.1_{\pm 1.9}$	$96.0_{\pm 0.3}$
p11	$84.7_{\pm 2.2}$	$92.7_{\pm 0.5}$	$91.4_{\pm 1.3}$	$94.1_{\pm 0.7}$
p12	$89.1_{\pm 1.2}$	$93.8_{\pm 0.3}$	$94.8_{\pm 2.0}$	$95.9_{\pm 1.4}$
p13	$89.1_{\pm 3.3}$	$91.0_{\pm 0.7}$	$93.2_{\pm 0.9}$	$95.6_{\pm 0.2}$
p14	$73.6_{\pm 2.0}$	$86.8_{\pm 2.0}$	$73.1_{\pm 2.4}$	$85.7_{\pm 2.4}$
p15	$81.6_{\pm 1.3}$	$83.3_{\pm 0.5}$	$75.6_{\pm 2.4}$	$84.8_{\pm 1.9}$
Average	$89.8_{\pm 0.5}$	$93.8_{\pm 0.3}$	$91.9_{\pm 0.7}$	$94.7_{\pm 0.4}$

1.2 More Visualizations

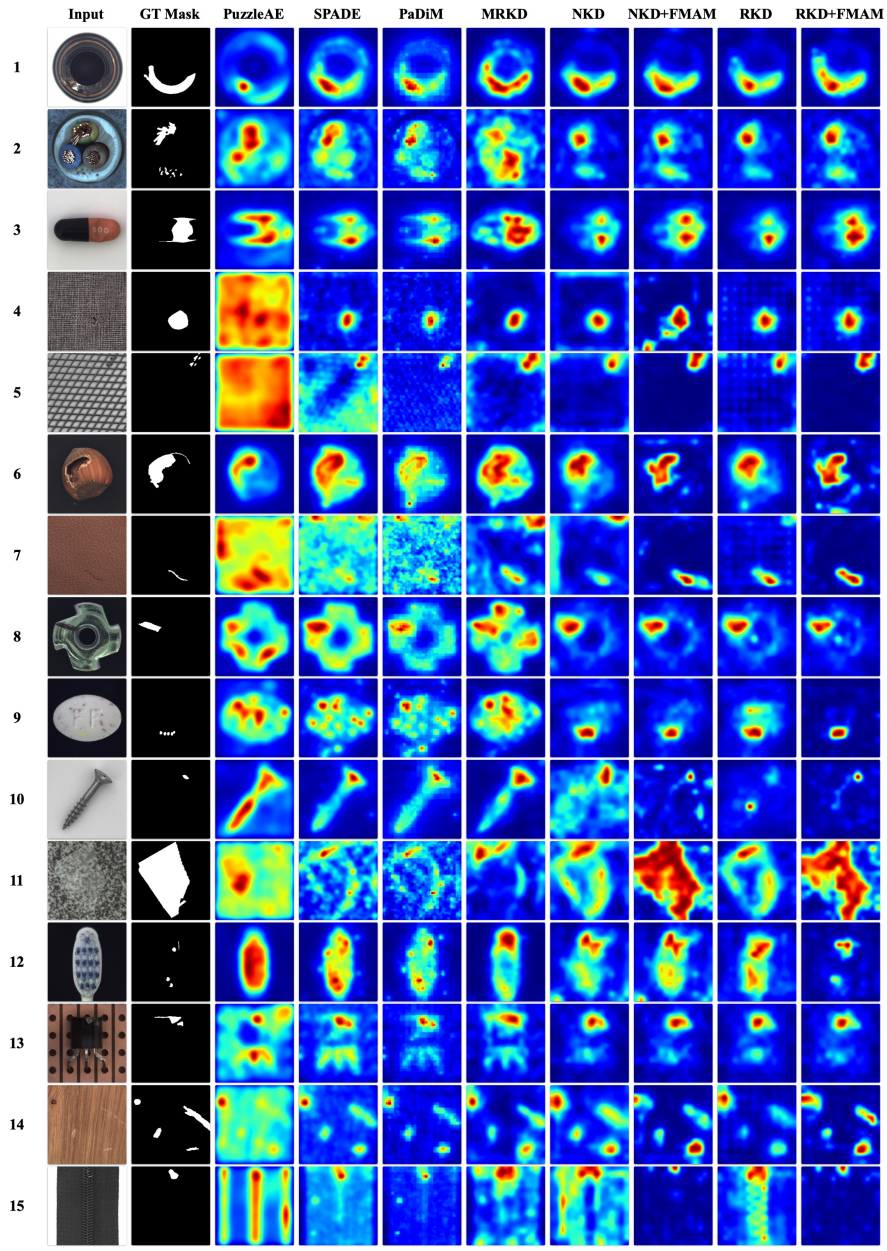

Figure 1: Visualization of anomaly detection and localization on the MVTec dataset. The first two columns show the original images and ground truth masks. The proposed feature mapping approach produces detections with less noise and higher accuracy compared to the original images.

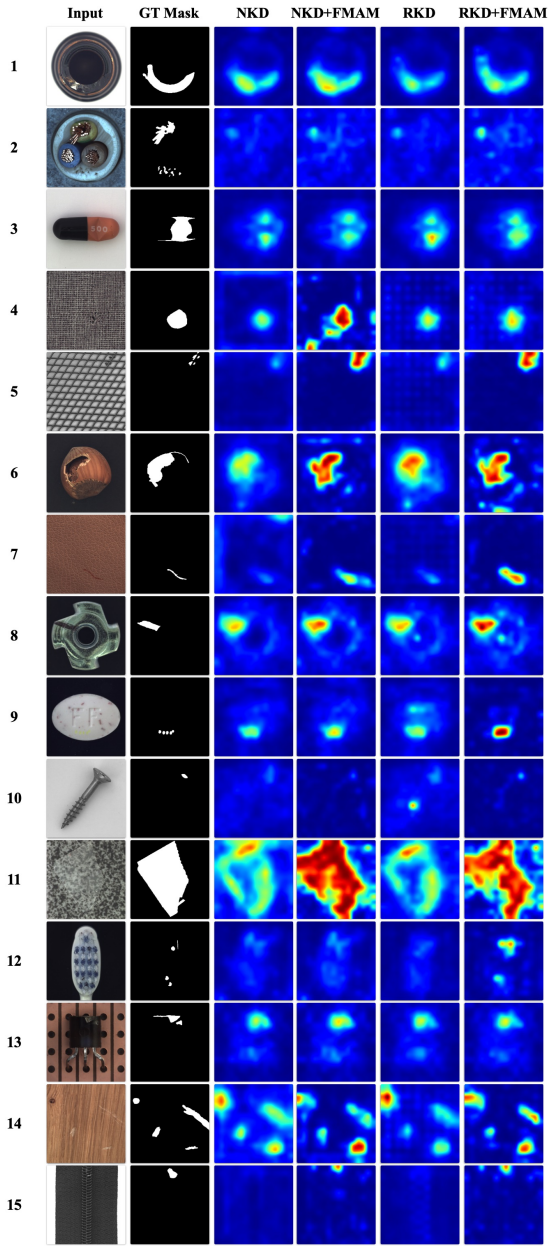

Figure 2: The results are normalized over all results for each product, rather than for each single heatmap as in Figure 1. This provides a global view of detection quality. The visualization shows that the proposed feature mapping successfully helps most MVTec products achieve a larger discrepancy between Teacher-Student models (rows 5, 6 ,7, 8, 9, 11, 12, 14 and 15).

1.3 Different Implementations of Feature Loss

There exists an inconsistency between the implementation of RKD and its original paper. Their paper claimed to use a vector-wise loss as the training loss for the T-S models, which calculates the feature discrepancy of T-S models at each point of the feature maps. However, their code implementation adopted a layer-wise loss, which calculates the feature discrepancy based on the entire feature map of a layer. This concern has been raised in the following link. Our observations indicate that this inconsistency can indeed affect the performance of anomaly detection.

To provide a comprehensive evaluation of these two implementations and different backbone models, we present the results in Table 5. The findings convincingly demonstrate that our proposed method consistently outperforms the various variations of RKD. Furthermore, we have included two additional performance metrics in Table 5: Per-Region-Overlap (PRO) and False Positive Rate (FPR) at a 95% True Positive Rate (TPR) threshold, to enhance the clarity and effectiveness of our method.

Table 5: Different RKD implementations of feature loss. (FPR=FPR@95%TPR)

Backbone	FMAM	Vector-wise loss				Layer-wise loss			
		IAUC	PAUC	PRO	FPR↓	IAUC	PAUC	PRO	FPR↓
Res18	✗	96.4	97.2	92.7	10.4%	97.8	97.1	91.3	11.1%
	✓	**98.2**	**97.3**	**93.2**	**10.2%**	**98.4**	**97.2**	**92.5**	**10.8%**
Res34	✗	96.4	97.3	92.7	9.4%	98.3	97.2	92.1	10.0%
	✓	**98.1**	**97.5**	**93.2**	9.4%	98.3	**97.4**	**92.5**	**9.8%**
Res50	✗	93.8	97.6	93.1	**8.9%**	98.5	97.6	93.1	8.3%
	✓	**95.8**	**97.7**	**93.4**	9.2%	**98.8**	**97.9**	**93.8**	**8.1%**
WRes50	✗	93.6	97.7	**93.6**	**8.8%**	98.5	97.7	94.0	7.6%
	✓	**96.1**	97.7	93.5	8.9%	**99.1**	**98.1**	**94.3**	**7.3%**