

Supplementary: FIND: Few-Shot Anomaly Inspection with Normal-Only Multi-Modal Data

Yiting Li¹, Fayao Liu^{1†}, Jingyi Liao¹, Sichao Tian², Chuan-Sheng Foo¹, Xulei Yang^{1†}

¹Institute for Infocomm Research, A*STAR, Singapore ²Chinese Academy of Medical Sciences and Peking Union Medical College, China

{li_yiting, liu_fayao, liao_jingyi, foo_chuan_sheng, yang_xulei}@i2r.a-star.edu.sg, {sctian}@implad.ac.cn

1. Eyecandies full-shot results

In the normal-only full-shot setting, FIND achieves the highest overall image level AUROC (I-AUROC) and AUPRO@30% among all methods. Specifically, FIND achieves an I-AUROC of 95.5% and AUPRO@30% of 92.5%, outperforming previous state-of-the-art approaches such as CFM, M3DM and 3D-ADNAS on average. For example, the next-best method (3D-ADNAS) reaches about 94.6% I-AUROC on Eyecandies. The performance gap is even larger compared to other baselines (CFM, M3DM), highlighting the superior anomaly detection and localization capabilities of FIND. The high mean scores indicate that FIND provides the most reliable overall performance on Eyecandies in terms of both detecting anomalous images and localizing anomalous regions.

2. Generating surface normals from point cloud

For MVTec-3D-AD, each sample provides an RGB image and a per-pixel XYZ map (a 3-channel TIFF file). We directly load this grid and estimate surface normals for every valid pixel using Open3D’s PCA-based local plane fitting ($k = 30$), followed by normal-orientation smoothing ($k = 50$). Pixels with zero coordinates or NaNs are left as zeros in the normal map. Following CFM and M3DM, we ignore background pixels. Both RGB and normal maps are resized to $224 * 224$ before being fed to the network.

3. Failure Cases

Although our approach benefits from the complementary knowledge of color and geometry, certain anomalies remain undetectable when both modalities struggle to provide clear discriminative cues. As shown in Fig. 1, these failures often arise when the defect is extremely subtle, such as slight surface scratches, which are overshadowed

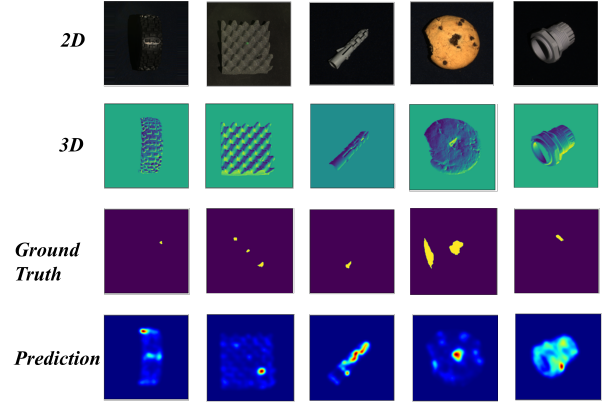


Figure 1. Visualization of failure cases.

by complex background patterns(e.g., object’s normal textures). Consequently, the fused anomaly map misidentifies the defect region, leading to false negatives or incomplete localization performance.

[†]Corresponding authors.

Table 1. Comparisons of per-category anomaly detection performance on Eyecandies.

Method	Candy	Choc. Cookie	Choc. Praline	Confetto	Gummy	Haz. Truffle	Lic. Sand.	Lollipop	Marsh.	Pep. Candy	Mean
I-AUROC											
EasyNet	73.7	93.4	86.6	96.6	71.7	82.2	84.7	86.3	97.7	96.0	86.9
AST	57.4	74.7	74.7	88.9	59.6	61.7	81.6	84.1	98.7	98.7	78.0
M3DM	62.4	95.8	95.8	100.0	88.6	78.5	94.9	83.6	100.0	100.0	89.7
CFM	68.0	93.1	95.2	88.0	86.5	78.2	91.7	84.0	99.8	96.2	88.1
3D-ADNAS	89.6	100.0	97.0	100.0	82.7	88.2	93.1	95.0	100.0	100.0	94.6
LSFA	67.0	95.4	96.1	100.0	91.3	76.7	94.3	85.4	100.0	100.0	90.6
FIND	83.8	100.0	96.2	100.0	95.2	88.8	95.3	96.7	100.0	99.0	95.5
AUPRO@30%											
AST	51.4	83.5	71.4	90.5	58.7	59.0	73.6	76.9	91.8	87.8	74.4
M3DM	90.6	92.3	80.3	98.3	85.5	68.8	88.0	90.6	96.6	95.5	88.2
CFM	94.2	90.2	83.1	96.5	87.5	76.2	79.1	91.3	93.9	94.9	88.7
3D-ADNAS	94.5	89.1	82.7	95.8	85.7	74.8	91.1	90.7	96.4	97.2	89.8
FIND	95.6	94.4	87.2	93.1	90.9	79.5	95.8	92.7	98.3	97.5	92.5