3CAD: A Large-Scale Real-World 3C Product Dataset for Unsupervised Anomaly Detection

Enquan Yang[1*], Peng Xing[2*], Hanyang Sun[1], Wenbo Guo[1], Yuanwei Ma[3], Zechao Li[2], Dan Zeng[1†]

[1]School of Communication and Information Engineering, Shanghai University
[2]Nanjing University of Science and Technology
[3]Changzhou Microintelligence Corporation
{enquanyang,sunhanyang,guowenbo,dzeng}@shu.edu.cn, {xingp_ng,zechao.li}@njust.edu.cn, yuanweima@gmail.com

Abstract

Industrial anomaly detection achieves progress thanks to datasets such as MVTec-AD and VisA. However, they suffer from limitations in the number of defect samples, types of defects, and availability of real-world scenes. These constraints inhibit researchers from further exploring the performance of industrial detection with higher accuracy. To this end, we propose a new large-scale anomaly detection dataset called 3CAD, which is derived from real 3C production lines. Specifically, the proposed 3CAD includes eight different types of manufactured parts, totaling 27,039 high-resolution images labeled with pixel-level anomalies. The key features of 3CAD are that it covers anomalous regions of different sizes, multiple anomaly types, and the possibility of multiple anomalous regions and multiple anomaly types per anomaly image. This is the largest and first anomaly detection dataset dedicated to 3C product quality control for community exploration and development. Meanwhile, we introduce a simple yet effective framework for unsupervised anomaly detection: a **C**oarse-to-**F**ine detection paradigm with **R**ecovery **G**uidance (CFRG). To detect small defect anomalies, the proposed CFRG utilizes a coarse-to-fine detection paradigm. Specifically, we utilize a heterogeneous distillation model for coarse localization and then fine localization through a segmentation model. In addition, to better capture normal patterns, we introduce recovery features as guidance. Finally, we report the results of our CFRG framework and popular anomaly detection methods on the 3CAD dataset, demonstrating strong competitiveness and providing a highly challenging benchmark to promote the development of the anomaly detection field. Data and code are available:

Code — https://github.com/EnquanYang2022/3CAD.

Introduction

The rapid expansion of 3C product manufacturing has surpassed the capabilities of traditional manual quality inspection, underscoring the need for advanced algorithms like Image Anomaly Detection (IAD) (Liu et al. 2024b). Deep learning approaches have demonstrated significant effectiveness in identifying complex and subtle defects in industrial images, enhancing both detection accuracy and robustness (Cao et al. 2024; Zhang et al. 2023a; Gu et al. 2024).

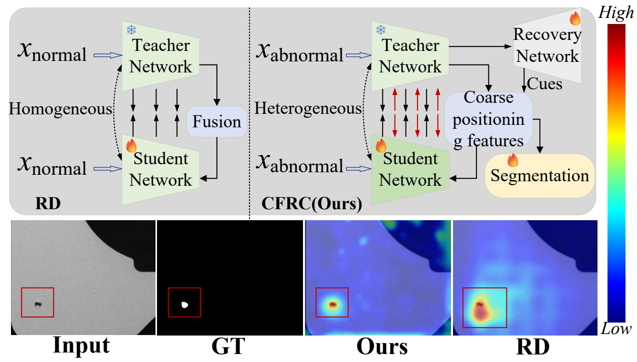

Figure 1: Comparison of previous anomaly detection distillation paradigm with our paradigm. First row: Left: reverse distillation; Right: our proposed paradigm.

Typically, supervised methods are employed when the types of defects are known and a sufficient amount of labeled data is available. These methods are based on classification (Nagata et al. 2019), detection (Zhang, Chang, and Jamshidi 2020), and segmentation models (Zou et al. 2018). However, obtaining defective samples and enough labeled data can be challenging in practice. As a result, unsupervised deep learning methods, which only require normal samples for training, have gained increasing attention (Roth et al. 2022; Liu et al. 2023; Cao et al. 2023; Liu et al. 2024a).

The introduction of widely-used industrial datasets such as MVTec-AD (Bergmann et al. 2019), VisA (Zou et al. 2022), and MPDD (Jezek et al. 2021a), has spurred significant academic interest in anomaly detection and led to numerous innovative approaches (Xing et al. 2023, 2024a). However, as research advances, several limitations of these datasets have become apparent. Firstly, the number of anomaly data poses a challenge. For instance, some categories of anomaly images in MVTec-AD consist of fewer than 60 samples, which may not adequately represent the performance of algorithms in practical applications. Secondly, there is a concern regarding the authenticity of anomalies. The defects in MVTec-AD, VisA, and Real-IAD (Wang et al. 2024) are primarily artificially created, such as scratches or deletions, despite being derived from real

---

[*]These authors contributed equally.
[†]Corresponding author

products. This creates a gap between the types of anomalies present in these datasets and those encountered in real-world production lines. Lastly, there is a tendency toward saturation in performance metrics. Recent methods achieve over 99% in both I-AUROC (image-level) and P-AUROC (pixel-level) metrics, making it challenging to discern the relative merits of new approaches. Additionally, there is currently no comprehensive anomaly detection dataset tailored specifically for 3C products, further highlighting a gap in the available resources.

To address the limitations of existing datasets and provide data closer to real scenes, we propose a new industrial anomaly detection dataset, named 3CAD, which is derived from real production lines. This dataset focuses on 3C product parts and offers the following advantages:

- Real-world Relevance. The dataset includes defects generated during the manufacturing process, reflecting real production scenarios with three common materials and eight different defect types.
- Large Scale. With 27,039 high-resolution images, it surpasses most existing anomaly detection datasets.
- Varied Defect Distribution. A single image may contain one or multiple defects of the same or different types, depending on the manufacturing process. The defect location is random.
- Complex Defect Morphology. Defects vary significantly in shape, size, and appearance, with many are similar to normal product features.
- Challenging Detection. Tiny and hidden defects make accurate detection difficult.

We assess several unsupervised anomaly detection algorithms using the 3CAD dataset. The results indicate that, although existing methods perform effectively on popular datasets, they encounter difficulties with precise defect localization in our dataset. This suggests there is significant potential for further improvement, as illustrated in Fig. 1.

To address the challenges of the 3CAD dataset, we propose a **C**oarse-to-**F**ine localization paradigm with **R**ecovery **G**uidance (CFRG), which enhances distillation and segmentation techniques by integrating a recovery task. CFRG consists of: Heterogeneous Feature Extraction: Teacher-student networks with different architectures extract diverse features from the same data, effectively pinpointing abnormal areas and addressing feature redundancy, especially for small and subtle defects. Recovery Network: Restores normal features from abnormal ones, capturing the underlying patterns of normal images. Segmentation Fine Localization: Leverages the recovery network's weights to guide the weighting of abnormal features extracted through distillation, which are then input into the segmentation network to improve localization accuracy. We hope this work advances anomaly detection in 3C products and inspires further research.

Related Work

Anomaly Detection Datasets: Datasets are crucial for defect detection research. Traditionally, algorithms are developed using specialized datasets for specific objects, such as pcbs (Tang et al. 2019), tiles (Huang, Qiu, and Yuan 2020), and steel (He et al. 2019). These training datasets often required manual labeling, limiting their impact on advancing industrial anomaly detection (IAD). The release of MVTec-AD in 2019 is a significant milestone, as it supported the development of unsupervised IAD algorithms by providing a diverse dataset. Subsequently, datasets like BTAD (Mishra et al. 2021), MPDD (Jezek et al. 2021b), and VisA (Zou et al. 2022) have further propelled IAD research. Recently, the Real-IAD (Wang et al. 2024) dataset introduced a larger, multi-view dataset, presenting new challenges for IAD.

Anomaly synthetic: Artificially synthesizing anomalies allows researchers to augment datasets and improve model performance, even with limited original data. Recent unsupervised anomaly detection methods have increasingly utilized synthetic anomaly images. For example, CutPaste (Li et al. 2021) generates negative samples by cutting and pasting image segments, while DRAEM (Zavrtanik, Kristan, and Skočaj 2021a) and DeSTSeg (Zhang et al. 2023b) use Perlin noise and random uniform samples to create anomaly masks. Additionally, DTD (Cimpoi et al. 2014) provides a source for blending anomalies into original images. The capabilities of diffusion models have further expanded synthetic data generation (Hu et al. 2024; Zhang, Xu, and Zhou 2024). However, synthetic approaches may yield unrealistic anomalies, and their diversity is limited by the inherent cognitive scope of the model.

Anomaly Detection Recovery-based methods train networks to restore defects in images to their normal state (Zavrtanik, Kristan, and Skočaj 2021b; Xing and Li 2023; Xing et al. 2024b). For example, RealNet (Zhang, Xu, and Zhou 2024) employs a feature reconstruction network with pre-trained multi-scale features, adaptively selecting and reconstructing residuals. By avoiding equal inputs and outputs, these methods mitigate the identity mapping issue in traditional reconstruction approaches. Moreover, the adaptability of diffusion models to various downstream tasks has spurred advancements in anomaly detection (Shen et al. 2023; He et al. 2024).

Knowledge distillation (KD) methods align teacher-student outputs for normal regions while differentiating defective ones for precise localization (Salehi et al. 2021). Identical network structures, however, may reduce feature diversity. Techniques like RD (Deng and Li 2022) and AST (Rudolph et al. 2023) address this by adopting serial or asymmetric architectures, improving differentiation between normal and abnormal features. Our method enhances this further by using heterogeneous teacher-student networks to better separate normal and abnormal features.

The 3CAD Dataset

Developing new datasets is crucial for maintaining high quality in 3C products and enhancing the effectiveness of anomaly detection for complex anomalies. Given the growing importance of unsupervised anomaly detection in industrial inspections, we introduce the 3CAD dataset, tailored for real-world 3C manufacturing scenarios.

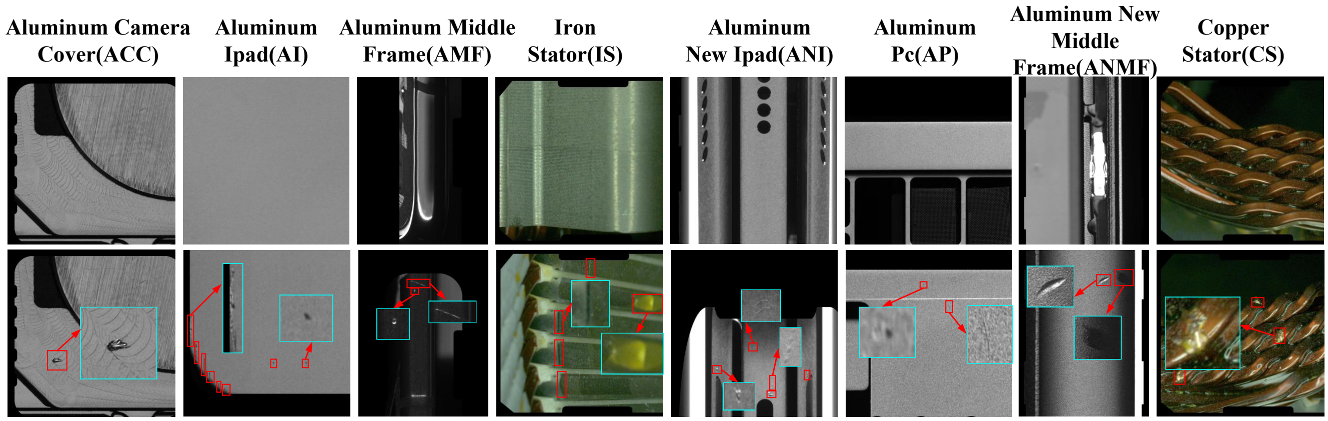

Figure 2: 3CAD dataset samples. The first row shows normal images, while the second row displays defective images.

Category	Training Images	Test Images(all)	Test Images(good)	Test Images(defect)	Defect types	Image Height	Image Width	NE / TE
ACC	784	1446	369	1077	10	288∼1024	288∼1024	1∼6/1∼1
AI	2096	2047	913	1134	3	760∼1024	600∼1024	1∼10/1∼2
AMF	1548	1479	731	748	5	540∼1024	800∼950	1∼9/1∼4
ANMF	1072	1406	670	736	6	400∼1024	430∼1024	1∼6/1∼2
ANI	2233	4936	999	3937	4	420∼1024	580∼1024	1∼23/1∼2
AP	1698	3161	911	2250	14	430∼1024	409∼1024	1∼12/1∼3
CS	409	959	196	763	1	1024∼1024	1024∼1024	1∼9/1∼1
IS	653	1112	295	817	4	1024∼1024	1024∼1024	1∼12/1∼2
All	10493	16546	5084	11462	47	-	-	-

Table 1: Statistical overview of the 3CAD dataset. The NE and TE in the last column indicate the number of anomalous regions and the number of anomalous types present in each defective image, respectively.

Data Construction

Data Source. The data originates from high-quality segmentation datasets accumulated by the company over several years from various production line projects. It is specifically tailored for defect detection in 3C products within industrial production processes.

Data Collection, Annotation. The data acquisition process consists of two stages. In the first stage, a series of dedicated mechanisms are employed, including loading, inspection, model analysis, and unloading. The workpiece is introduced into this automated system, where images of each part are captured. These images are then analyzed by a model to detect defects. During the commissioning stage, specialized quality control staff collect both defective and non-defective materials from the production line and photograph them using the designated equipment. Subsequently, labeling staff annotate the images with pixel-level precision based on the assessments of Quality Engineers (QEs), using a self-developed labeling software similar to LabelMe.

Data Cleaning. Data cleaning and labeling are seamlessly integrated to ensure high-quality results. The algorithm autonomously handles certain cleaning tasks, such as consolidating data types based on its schema. To maintain high standards in labeling, annotators undergo daily evaluations to ensure their understanding of defect morphology is accurate and consistent.

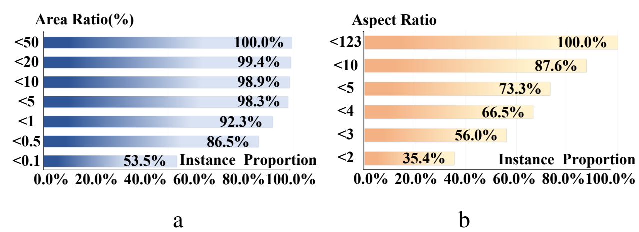

Figure 3: Statistics of the proposed 3CAD dataset: a) Defect area ratio. b) Aspect ratio of the minimum bounding rectangle for the defect area.

Additionally, a review system is in place where the labeling team leader assesses the work of annotators. If any inconsistencies or quality issues are detected, the data is sent back for re-labeling. Typically, under standard project conditions and without interruptions, each annotator processes an average of 10 to 50 samples per day. For details on dataset construction, see the supplementary material.

Dataset Description. The details of the proposed 3CAD dataset are shown in Tab. 1. (1) It comprises 10,493 training images and 16,546 testing images that are carefully selected to represent the best acquisition for each product type. Each defect image is labeled with high-quality pixel-level defect labels. (2) 3CAD covers a wide range of 3C products from real-world product lines, including complex items like the

Dataset	Normal	Abnormal	Defect Source	MDI	MDT
BTAD	2250	580	Real	✗	✗
MPDD	1064	282	Real	✓	✗
MVTec-AD	4096	1258	Forged	✗	✗
VisA	10621	1200	Forged	✓	✗
Real-IAD	99721	51329	Forged	✗	✗
Ours	15577	11462	Real	✓	✓

Table 2: Comparison with the current popular IAD datasets. ✓: Satisfied. ✗: Not satisfied. MDI: Multiple Defect Instances in One Image; MDT: Multiple Defect Types in One Image.

Aluminum Camera and simple items like Ipad product. (3) Rich anomaly categories. Defect types in the dataset range from prominent ones like bumps and dents to more subtle ones such as scratches and pinholes. Meanwhile, a single defect image may simultaneously contain multiple types or multiple abnormal regions (as shown in Fig. 2). (4) Diverse anomalous regions. As shown in Fig. 3.a, 3CAD covers both large-scale anomalous regions and small-scale anomalous areas, with a greater proportion of the currently challenging small-scale anomalies. Fig 3.b depicts the distribution of the number of minimum outer rectangular aspect ratios of the anomalous regions, showing the diversity of the morphological distribution of the anomalous regions in 3CAD.

Comparison with Popular Datasets. Tab. 2 shows the results of comparing the proposed 3CAD with other popular datasets. Firstly, 3CAD is derived directly from real manufacturing environments, capturing diverse scenarios where the same defects can exhibit significant variation and different defects can appear quite similar. In contrast, datasets such as MVTec-AD, VisA, and Real-IAD tend to rely on artificially generated defect types and distributions that are not conducive to real-world applications. The authenticity of 3CAD ensures that algorithms trained on it are better suited for real-world deployment. Moreover, 3CAD includes 15,577 normal samples and 11,462 abnormal samples, surpassing BTAD and MPDD in both scale and variety of defect types. Last but not least, 3CAD is unique in its may inclusion of multiple defect instances and types in single image. While other datasets like MPDD and VisA allow for multiple instances, they contain a limited number of defect instances and they fall short in presenting multiple defect types within a single image. This feature of 3CAD is particularly important as it mirrors real-world scenarios where products may have more defect instances and more than one type of defect occurring simultaneously, challenging the detection algorithms to identify each one accurately.

Method

To tackle the challenges posed by 3CAD, we propose a simple but effective framework, **C**oarse-to-**F**ine detection paradigm with **R**ecovery **G**uidance, called CFRG, whose framework is shown in Fig. 4. To mitigate the challenges posed by small defects in 3CAD and the discrepancy between real and synthetic anomalies, we propose to use distillation for initial coarse localization, followed by guidance through an anomaly recovery task, and finally fine localization of the anomalies through a segmentation network.

Knowledge Distillation for Coarse Localization

To localize anomalous regions, we use knowledge distillation with a pre-trained teacher network (WideResNet50 (Zagoruyko and Komodakis 2016)) and a learnable student network (EfficientNet-b0 (Tan and Le 2019)). The student mimics the behavior of the teacher on normal samples and learns to distinguish anomalies, reducing mislocalization, especially for subtle defects. The heterogeneous design, where the teacher and student focus on different aspects of feature extraction, enhances the detection of small, weak, and background-like anomalies.

Given an input image $x_n \in \mathbb{R}^{C\times H\times W}$, anomalies are synthesized using DTD and 2D Perlin noise to produce x_a. Multi-level features $\{F_i^{t_a}\}_{i=1}^{K} \in \mathbb{R}^{C_i\times H_i\times W_i}$ and $\{F_i^{s_a}\}_{i=1}^{K} \in \mathbb{R}^{C_i\times H_i\times W_i}$ are extracted from x_a by teacher and student networks. We start by calculating the cosine similarity at each stage:

$$\mathcal{L}_i^{\cos}(x,y) = 1 - \frac{F_i^{t_a}(x,y)}{\|F_i^{t_a}(x,y)\|_2} \cdot \frac{F_i^{S_a}(x,y)}{\|F_i^{S_a}(x,y)\|_2}, \quad (1)$$

where indices i and j denote the spatial coordinates on the feature map. Next, using the ground-truth mask to differentiate between normal and abnormal regions, we construct a loss function that maximizes the cosine similarity within normal regions while minimizing it in abnormal regions:

$$L_{dis} = \sum_{i=1}^{3}\left(\left(1-G\right)\mathcal{L}_i^{\cos}(x,y) + G\left(1-\mathcal{L}_i^{\cos}(x,y)\right)\right), \quad (2)$$

where $G \in \{0,1\}$, indicating the Ground Truth mask.

Recovery Feature as Guidance

Recovery-based methods train networks to treat anomalies as noise and focus on reconstructing images to their normal state. This approach helps the network learn and represent the intrinsic patterns of normal images, reducing sensitivity to real-world anomalies and mitigating distillation localization bias, even when simulated and real anomalies differ. In our framework, a pre-trained teacher network extracts multi-scale features $\{F_i^{t_n}\}_{i=1}^{K}$ and $\{F_i^{t_a}\}_{i=1}^{K}$ from normal and abnormal images, respectively. These features capture both global and local details, enhancing model robustness to noise. Following RD (Deng and Li 2022), we fuse $\{F_i^{t_a}\}_{i=1}^{K}$ into the dimension of the final layer and input it into a recovery network. This network, using a ResNet block with transposed convolution, generates multi-layer features $\{F_i^{r}\}_{i=1}^{K}$ that match the dimensions of $\{F_i^{t_a}\}_{i=1}^{K}$.

$$\{F_i^{r}\}_{i=1}^{K} = Resblock\left(Bn\left(F_i^{t_a}\right)\right), \quad (3)$$

where Bn denotes the feature fusion operation, while $Resblock$ refers to the stacked ResNet blocks. Our goal is to align the feature spaces of $\{F_i^{t_n}\}_{i=1}^{K}$ and $\{F_i^{r}\}_{i=1}^{K}$ to optimize the recovery network. We still use cosine similarity loss to achieve this to obtain L_{rec}.

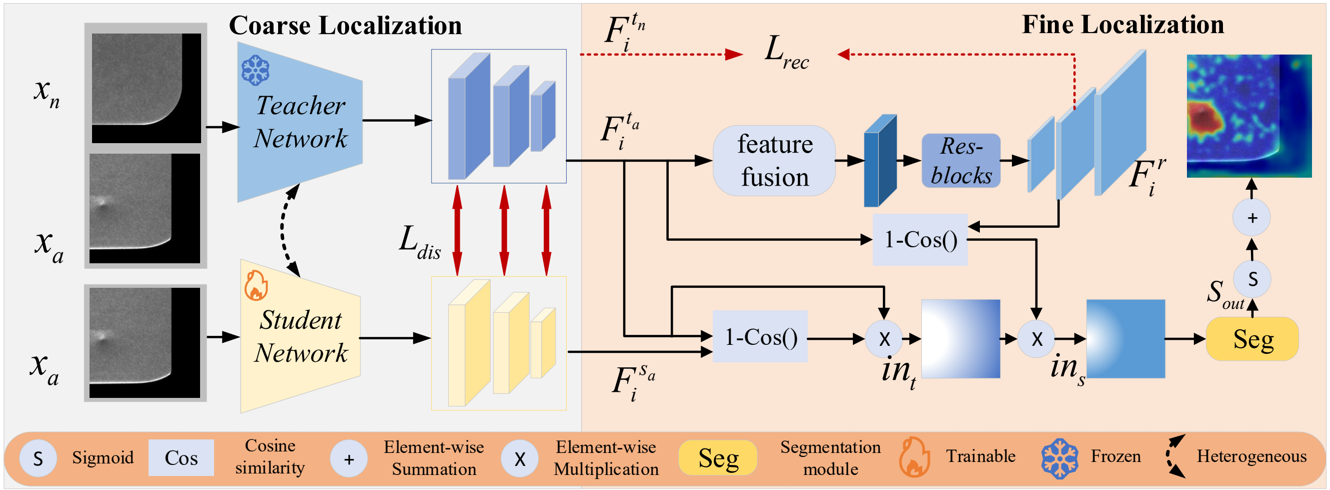

Figure 4: The proposed CFRG framework comprises two components: 1) a distilled localization network and 2) a refined segmentation network with restored hints. During training, in the first stage, x_a and x_n are input into the teacher network, while x_a is input into the student network, and the distillation loss between the teacher and the student is calculated. In the second stage, the teacher's features are weighted using the first-stage localization weights and the recovery branch's hint weights, then input into the segmentation network. During testing, the recovery branch generates the localization result from the input and $\{F_i^r\}_{i=1}^K$, which is then added to the output S_{out} of the segmentation network to obtain the final anomaly map.

Segmentation for Fine-Grained Localization

Pixel-level localization by distillation and recovery network works well for clear anomalies but is insufficient for more complex cases in our data. To improve precision, we integrate segmentation techniques.

First, we compute the cosine similarity w_d between $\{F_i^{t_a}\}_{i=1}^K$ and $\{F_i^{s_a}\}_{i=1}^K$, where lower values indicate abnormal regions. The anomaly localization is represented by $1 - w_d$, applied to the features of teacher network to obtain in_t. A similar process is applied to the input and output of recovery network, yielding $1 - w_r$ to refine in_t and produce in_s. Finally, in_s is input into a segmentation module, which fuses multi-layer information through a skip connection and outputs the segmentation mask S_{out} in the original image size. The segmentation module is optimized using binary cross-entropy loss.

$$L_{bce} = \frac{1}{H \times W} \sum_{x=1} \sum_{y=1} [-G(x,y) \log S_{out}(x,y) - (1 - G(x,y) \log(1 - S_{out}(x,y))], \tag{4}$$

Total loss of CFRG is expressed as:

$$L_{all} = L_{dis} + L_{rec} + L_{bce} \tag{5}$$

Model Inference

During the inference phase, we measure the cosine similarity between the output of teacher and$\{F_i^r\}_{i=1}^K$. This result is then combined with the segmentation mask S_{out} to obtain the final anomaly score map. We apply Gaussian filtering with $\sigma = 4$ to achieve smooth boundaries.

Experiments

Experimental Settings

Datasets. We conduct benchmark experiments on the 3CAD and MVTec-AD datasets. The MVTec-AD dataset consists of 5,354 high-resolution images from various domains, covering 5 textures and 10 objects. The training set includes 3,629 normal images, while the test set comprises 1,725 images.

Evaluation Metrics. Following existing methodologies, we use Area Under the Receiver Operator Curve (AUROC) and Pixel-wise Per-Region Overlap (P-PRO) for anomaly detection and localization. Specifically, I-AUROC represents image-level anomaly detection, and P-AUROC represents pixel-level anomaly localization. Additionally, we adopt Pixel-wise Average Precision (AP) to further compare the performance in detecting abnormal regions.

Implemental Details. All images are resized to 256 × 256. During training, we use the AdamW optimizer with a learning rate of 0.0005. The learning rate decays by a factor of 0.2 at epochs 40 and 45. We train for 50 epochs on a single NVIDIA RTX 3090 24GB with a batch size of 32.

Benchmark Evaluation

Baseline Methods. For benchmarking and performance comparison, we select embedding-based (E-b) methods PaDiM (Defard et al. 2021), FastFlow (Yu et al. 2021), RD (Deng and Li 2022), SimpleNet (Liu et al. 2023), RD++ (Tien et al. 2023); synthesis-based (A-syn) methods DREAM (Zavrtanik, Kristan, and Skočaj 2021a), DeSTSeg (Zhang et al. 2023b); reconstruction-based (R-b) method RealNet (Zhang, Xu, and Zhou 2024); and unified (U-ni)

	Method	ACC	AI	AMF	ANMF	ANI	AP	CS	IS	Mean
E-b	**PaDiM**	85.5/-	93.9/-	92.6/-	85.1/-	87.6/-	79.8/-	87.8/-	76.9/-	86.1/-
	FastFlow	77.1/-	82.8/-	68.5/-	81.2/-	78.5/-	52.8/-	71.2/-	70.7/-	72.8/-
	RD	92.2/34.3	96.8/8.0	97.2/3.9	92.5/1.9	94.7/9.4	87.4/1.8	93.1/5.7	84.8/4.7	92.3/8.7
	RD++	91.1/32.9	96.6/5.1	97.4/5.3	91.6/2.0	95.4/12.5	85.5/1.6	93.1/6.8	84.9/4.6	91.9/8.8
	SimpleNet	75.9/13.3	95.4/13.0	93.8/5.5	69.6/0.4	93.1/9.9	66.7/0.7	83.4/1.9	81.5/5.3	82.4/6.3
A-syn	**DREAM**	63.5/20.3	94.8/20.4	92.3/4.5	74.6/1.8	81.4/15.6	69.5/1.2	91.0/6.2	76.9/4.9	80.5/9.4
	DeSTSeg	87.8/32.5	96.9/12.5	96.6/4.1	94.8/5.8	93.2/9.2	77.1/2.2	90.8/3.1	**86.9**/8.1	90.5/9.6
R-b	**RealNet**	81.0/-	92.4/-	89.3/-	82.9/-	89.8/-	76.0/-	81.1/-	78.3/-	83.8/-
U-ni	**UniAD**	84.7/-	94.6/-	93.5/-	88.0/-	86.0/-	81.4/-	85.1/-	80.9/-	86.8/-
	CRAD	**92.9**/-	96.7/-	97.0/-	90.8/-	92.8/-	**88.4**/-	91.0/-	86.4/-	92.0/-
E-b	**Ours**	91.1/**34.6**	**97.5/24.6**	**98.3/23.3**	**95.9/13.1**	**96.9/23.1**	88.2/**2.5**	**93.5/10.2**	85.9/**9.3**	**93.4/17.6**

Table 3: Performance of popular IAD algorithms and our paradigm on 3CAD. We report the P-AUROC (%) and AP (%) metrics for each class, along with the average across all classes. Higher values indicate better performance.

	Method	ACC	AI	AMF	ANMF	ANI	AP	CS	IS	Mean
E-b	**PaDiM**	81.6/-	**96.1/-**	89.3/-	67.1/-	79.1/-	79.5/-	66.4/-	63.5/-	77.8/-
	FastFlow	79.1/-	89.5/-	82.3/-	63.0/-	71.9/-	71.3/-	63.4/-	64.3/-	73.1/-
	RD	90.6/82.7	96.0/81.5	89.5/82.1	66.8/70.0	81.8/81.3	79.0/72.1	74.0/77.1	65.9/60.8	80.4/75.9
	RD++	92.0/83.2/	95.5/86.5	89.0/84.0	70.0/65.7	81.8/85.2	80.5/70.8	76.3/**78.7**	67.7/**63.8**	81.6/77.2
	SimpleNet	81.6/54.8	92.9/70.4	85.1/73.0	61.0/40.7	76.9/64.7	69.5/55.6	71.5/55.3	67.2/54.6	75.7/58.7
A-syn	**DREAM**	80.4/-	89.4/-	73.3/-	61.9/-	78.5/-	71.7/-	68.1/-	70.8/-	74.3/-
	DeSTSeg	91.9/79.7	95.1/93.5	93.1/85.9	77.3/81.3	87.5/75.1	85.6/77.7	70.4/66.5	**86.3**/60.0	85.9/77.4
R-b	**RealNet**	83.9/43.3	90.7/70.4	73.9/38.1	66.6/27.7	70.0/22.2	70.4/40.1	65.2/47.6	64.3/13.1	73.1/37.8
U-ni	**UniAD**	82.4	93.3/-	87.4/-	65.5/-	86.4/-	72.4/-	56.8/-	62.4/-	75.8/-
	CRAD	88.2/-	92.6/-	89.4/-	69.0/-	75.6/-	82.7/-	72.6/-	64.2/-	79.3/-
E-b	**Ours**	**93.9/84.6**	**96.1/91.8**	**94.5/90.6**	**83.8/82.4**	**90.7/88.4**	**87.2/78.5**	**77.2**/78.5	68.4/61.3	**86.5/82.0**

Table 4: Performance of popular IAD algorithms and our paradigm on 3CAD. We report the I-AUROC (%) and P-PRO (%) metrics for each class, along with the average across all classes. Higher values indicate better performance.

methods UniAD (You et al. 2022), CRAD (Lee et al. 2024). We use Anomalib to reproduce PaDiM and FastFlow, while the remaining methods are implemented using their official code with original configurations.

Results on 3CAD. The results of all methods on the 3CAD dataset are shown in Tab. 3 and 4. We find significant differences in performance among popular methods. Embedding-based methods, particularly teacher-student networks, perform well due to efficient feature extraction, embedding space construction, and similarity measurement, while flow models performs slightly behind. Despite the challenges posed by the deviation of synthetic anomalies from real ones, anomaly synthesis-based methods still achieve strong localization, particularly in I-AUROC and AP metrics. Combining anomaly synthesis with embedding-based approaches further improves performance. However, traditional reconstruction-based methods struggle with the challenges of the 3CAD dataset, and unified anomaly detection methods still have considerable room for improvement due to the complex distribution of each subset.

CFRG on 3CAD. The proposed CFRG method is evaluated on the 3CAD dataset, with results shown in Tab.3 and 4. CFRG achieve 93.4% AUROC, 86.5% AUPRO, 82.0% AP, and 17.6%. Compared to embedding-based methods, CFRG improve P-AUROC by 1.1% over RD, I-AUROC by 4.9% over RD++, PRO by 4.8%, and AP by 8.8%. Against anomaly synthesis methods, CFRG outperform DeSTSeg in AP by 7.3%. For reconstruction-based methods, CFRG surpass RealNet in P-AUROC and I-AUROC by 9.6% and 13.3%, respectively. When compared to unified anomaly detection paradigms, CFRG outperform UniAD by 6.6% in P-AUROC and 10.7% in I-AUROC, and CRAD by 1.4% and 7.2%, respectively. These results confirm the effectiveness of its coarse-to-fine localization approach over stand-alone reconstruction and distillation methods.

Methods	MVTec-AD/3CAD			
	P-AUROC	I-AUROC	P-PRO	AP
FastFlow	98.5/72.8	99.4/73.1	-	-
RD++	98.2/91.9	99.4/81.6	94.977.2	61.5/8.8
SimpleNet	98.1/82.4	99.6/75.7	89.9/58.7	-
DREAM	97.3/80.5	98.0/74.3	-	68.4/9.4
DeSTSeg	97.9/90.5	98.6/85.9	94.7/77.4	75.8/9.6
RealNet	98.9/83.8	99.5/73.1	91.4/37.8	-
UniAD	96.8/86.8	96.5/75.8	-	-
CRAD	97.8/92.0	99.3/79.3	-	-
Ours	98.4/93.4	98.4/86.5	95.6/82.0	73.4/17.6

Table 5: Performance comparison of popular IAD algorithms on MVTec-AD, averaged across all categories. (-) indicates unavailable metrics in the official paper.

Comparison of Results. We compare the selected benchmarks on 3CAD and MVTec-AD, revealing several key observations, as shown in Tab. 5. Firstly, while CFRG performs well on 3CAD, its performance is significantly lower than

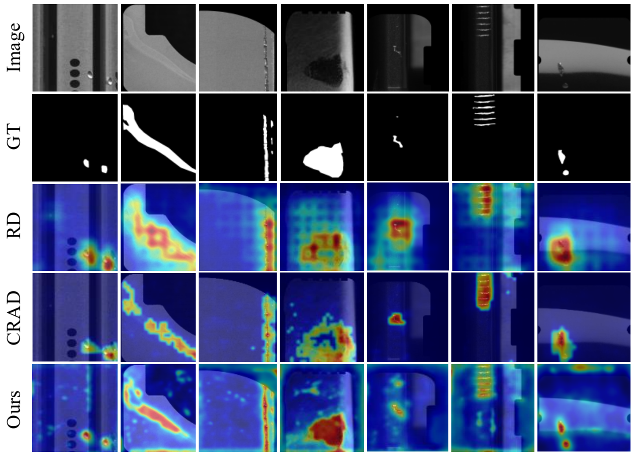

Figure 5: Qualitative illustration on 3CAD dataset.

on MVTec-AD. Secondly, most popular methods, which achieve around 99% on MVTec-AD, experience a performance drop of over 10% on our dataset. Unified-based methods also show a substantial decline due to the diversity and complexity of 3CAD, making it more challenging. Lastly, the unique complexity of 3CAD hinders existing methods, with AP metrics often below 10% even when AUROC is high. P-PRO varies between 60%-90%, underscoring the increased difficulty of fine pixel-level localization in 3CAD.

Qualitative Results. In Fig. 5, We conduct qualitative experiments on 3CAD. Our method accurately locates defects with large morphological spans, providing finer boundaries. In densely arranged thin scratches (third and sixth columns), RD and CRAD mislocate them. In the first, fifth, and seventh columns, our method effectively pinpoints small local defects, while RD and CRAD produce wide and fuzzy localization errors. For non-structural large-area defects (second and fourth columns), our method offers more comprehensive and precise coverage compared to RD and CRAD.

Ablation Studies

Study on Architecture Design of CFRG. We study the effectiveness of two key components: the recovery branch and the segmentation module. As shown in Tab. 6, WRC indicates the removal of the recovery branch. Retaining this branch allows the network to capture the underlying patterns of normal images when facing real anomalies, enhancing anomaly modeling. However, it may slightly interfere with the AP indicator. Overall, the combined effect is optimal, with improvements of 6.8% on P-AUROC, 4.3% on I-AUROC, and 22.2% on P-PRO. WS indicates the removal of the segmentation branch. The segmentation network's ability to process local features brings significant improvements, with increases of 2.8% on P-AUROC, 3.1% on I-AUROC, 7.2% on P-PRO, and 9.1% on AP.

Study on Distillation Paradigms. The feature distribution of the distillation process significantly impact localization performance, as shown in Tab. 6.WP denotes that we removed the practice of pushing the abnormal feature distance between the teacher and student networks, focusing solely on closing the normal feature distance. This led to decreases of 0.1%, 0.6%, 0.5%, and 1.1% across the four metrics. When dealing with weak defects, the introduction of this push-pull paradigm aids in separating abnormal areas. This approach aligns with our goal of using heterogeneous teacher and student networks.

WRC		✓	✓	✓	✓	✓
WS	✓		✓	✓	✓	✓
WP	✓	✓		✓	✓	✓
WC	✓	✓	✓		✓	✓
WH	✓	✓	✓	✓		✓
P-AUROC↑	86.6	90.6	93.3	93.0	92.9	**93.4**
I-AUROC↑	82.2	79.1	85.9	85.9	85.7	**86.5**
P-PRO↑	59.8	74.8	81.5	81.3	80.9	**82.0**
AP↑	**19.7**	8.5	16.5	16.1	17.1	17.6

Table 6: Ablation studies of CFRG with AUROC, P-PRO, and AP metrics.

Study on Recovery guidence. We further examined the impact of the hint weight from the recovery branch. In the WC scenario, we used only in_t as the input to the segmentation network, removing the auxiliary weight from the recovery branch. This led to performance drops of 0.3%, 0.6%, 0.7%, and 1.5% across the four metrics. This demonstrates that the recovery branch not only enhances the final prediction but also helps reduce distribution deviation introduced during the distillation stage.

Study on Heterogeneous. We use heterogeneous teacher and student networks to ensure that weak defect regions have distinct feature representations, promoting feature separation and accelerating network optimization during distillation. In the WH scenario, we replaced the student network with WideResNet50, the same architecture as the teacher. This change resulted in performance drops of 0.5%, 0.8%, 1.1%, and 0.5% across the four metrics. The improvement achieved by using a student network with a different architecture highlights the potential for further research within the heterogeneous paradigm.

Conclusion

Given the current state of industrial anomaly detection, this paper introduces 3CAD, the first large-scale dataset focused on anomaly detection in the manufacturing process of 3C product parts. Derived from real production lines, 3CAD provides pixel-level annotations for eight different types of industrial product parts, comprising 27,039 high-resolution images. The samples and defects offer a realistic representation of manufacturing scenarios. To tackle the challenges of 3CAD, we propose CFRG, a method that uses heterogeneous teacher-student networks to generate coarse localization, which is then refined through a segmentation network with recovery guidence. Our approach demonstrates excellent performance. Future work will focus on developing more advanced localization paradigms to further enhance accuracy.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China (Grant No. 62372284).

References

Bergmann, P.; Fauser, M.; Sattlegger, D.; and Steger, C. 2019. MVTec AD–A comprehensive real-world dataset for unsupervised anomaly detection. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*.

Cao, Y.; Xu, X.; Liu, Z.; and Shen, W. 2023. Collaborative discrepancy optimization for reliable image anomaly localization. *IEEE Transactions on Industrial Informatics*.

Cao, Y.; Xu, X.; Zhang, J.; Cheng, Y.; Huang, X.; Pang, G.; and Shen, W. 2024. A survey on visual anomaly detection: Challenge, approach, and prospect. *arXiv preprint arXiv:2401.16402*.

Cimpoi, M.; Maji, S.; Kokkinos, I.; Mohamed, S.; and Vedaldi, A. 2014. Describing textures in the wild. In *Proceedings of the IEEE conference on computer vision and pattern recognition*.

Defard, T.; Setkov, A.; Loesch, A.; and Audigier, R. 2021. Padim: a patch distribution modeling framework for anomaly detection and localization. In *International Conference on Pattern Recognition*. Springer.

Deng, H.; and Li, X. 2022. Anomaly detection via reverse distillation from one-class embedding. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*.

Gu, Z.; Zhang, J.; Liu, L.; Chen, X.; Peng, J.; Gan, Z.; Jiang, G.; Shu, A.; Wang, Y.; and Ma, L. 2024. Rethinking Reverse Distillation for Multi-Modal Anomaly Detection. In *Proceedings of the AAAI Conference on Artificial Intelligence*.

He, H.; Zhang, J.; Chen, H.; Chen, X.; Li, Z.; Chen, X.; Wang, Y.; Wang, C.; and Xie, L. 2024. A diffusion-based framework for multi-class anomaly detection. In *Proceedings of the AAAI Conference on Artificial Intelligence*.

He, Y.; Song, K.; Meng, Q.; and Yan, Y. 2019. An end-to-end steel surface defect detection approach via fusing multiple hierarchical features. *IEEE transactions on instrumentation and measurement*.

Hu, T.; Zhang, J.; Yi, R.; Du, Y.; Chen, X.; Liu, L.; Wang, Y.; and Wang, C. 2024. Anomalydiffusion: Few-shot anomaly image generation with diffusion model. In *Proceedings of the AAAI Conference on Artificial Intelligence*.

Huang, Y.; Qiu, C.; and Yuan, K. 2020. Surface defect saliency of magnetic tile. *The Visual Computer*.

Jezek, S.; Jonak, M.; Burget, R.; Dvorak, P.; and Skotak, M. 2021a. 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)*. IEEE.

Jezek, S.; Jonak, M.; Burget, R.; Dvorak, P.; and Skotak, M. 2021b. 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)*. IEEE.

Lee, J. C.; Kim, T.; Park, E.; Woo, S. S.; and Ko, J. H. 2024. Continuous Memory Representation for Anomaly Detection. In *European conference on computer vision*. Springer.

Li, C.-L.; Sohn, K.; Yoon, J.; and Pfister, T. 2021. Cutpaste: Self-supervised learning for anomaly detection and localization. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*.

Liu, J.; Wu, K.; Nie, Q.; Chen, Y.; Gao, B.-B.; Liu, Y.; Wang, J.; Wang, C.; and Zheng, F. 2024a. Unsupervised Continual Anomaly Detection with Contrastively-Learned Prompt. In *Proceedings of the AAAI Conference on Artificial Intelligence*.

Liu, J.; Xie, G.; Wang, J.; Li, S.; Wang, C.; Zheng, F.; and Jin, Y. 2024b. Deep industrial image anomaly detection: A survey. *Machine Intelligence Research*.

Liu, Z.; Zhou, Y.; Xu, Y.; and Wang, Z. 2023. Simplenet: A simple network for image anomaly detection and localization. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*.

Mishra, P.; Verk, R.; Fornasier, D.; Piciarelli, C.; and Foresti, G. L. 2021. VT-ADL: A vision transformer network for image anomaly detection and localization. In *2021 IEEE 30th International Symposium on Industrial Electronics (ISIE)*. IEEE.

Nagata, F.; Tokuno, K.; Nakashima, K.; Otsuka, A.; Ikeda, T.; Ochi, H.; Watanabe, K.; and Habib, M. K. 2019. Fusion method of convolutional neural network and support vector machine for high accuracy anomaly detection. In *2019 IEEE International Conference on Mechatronics and Automation (ICMA)*. IEEE.

Roth, K.; Pemula, L.; Zepeda, J.; Schölkopf, B.; Brox, T.; and Gehler, P. 2022. Towards total recall in industrial anomaly detection. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*.

Rudolph, M.; Wehrbein, T.; Rosenhahn, B.; and Wandt, B. 2023. Asymmetric student-teacher networks for industrial anomaly detection. In *Proceedings of the IEEE/CVF winter conference on applications of computer vision*.

Salehi, M.; Sadjadi, N.; Baselizadeh, S.; Rohban, M. H.; and Rabiee, H. R. 2021. Multiresolution knowledge distillation for anomaly detection. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*.

Shen, F.; Ye, H.; Zhang, J.; Wang, C.; Han, X.; and Yang, W. 2023. Advancing pose-guided image synthesis with progressive conditional diffusion models. *arXiv preprint arXiv:2310.06313*.

Tan, M.; and Le, Q. 2019. Efficientnet: Rethinking model scaling for convolutional neural networks. In *International conference on machine learning*. PMLR.

Tang, S.; He, F.; Huang, X.; and Yang, J. 2019. Online PCB defect detector on a new PCB defect dataset. *arXiv preprint arXiv:1902.06197*.

Tien, T. D.; Nguyen, A. T.; Tran, N. H.; Huy, T. D.; Duong, S.; Nguyen, C. D. T.; and Truong, S. Q. 2023. Revisiting reverse distillation for anomaly detection. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*.

Wang, C.; Zhu, W.; Gao, B.-B.; Gan, Z.; Zhang, J.; Gu, Z.; Qian, S.; Chen, M.; and Ma, L. 2024. 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*.

Xing, P.; and Li, Z. 2023. Visual anomaly detection via partition memory bank module and error estimation. *IEEE Transactions on Circuits and Systems for Video Technology*.

Xing, P.; Sun, Y.; Zeng, D.; and Li, Z. 2023. Normal image guided segmentation framework for unsupervised anomaly detection. *IEEE Transactions on Circuits and Systems for Video Technology*.

Xing, P.; Tang, H.; Tang, J.; and Li, Z. 2024a. ADPS: Asymmetric Distillation Postsegmentation for Image Anomaly Detection. *IEEE Transactions on Neural Networks and Learning Systems*.

Xing, P.; Zhang, D.; Tang, J.; et al. 2024b. A Recover-then-Discriminate Framework for Robust Anomaly Detection. *arXiv preprint arXiv:2406.04608*.

You, Z.; Cui, L.; Shen, Y.; Yang, K.; Lu, X.; Zheng, Y.; and Le, X. 2022. A unified model for multi-class anomaly detection. *Advances in Neural Information Processing Systems*.

Yu, J.; Zheng, Y.; Wang, X.; Li, W.; Wu, Y.; Zhao, R.; and Wu, L. 2021. Fastflow: Unsupervised anomaly detection and localization via 2d normalizing flows. *arXiv preprint arXiv:2111.07677*.

Zagoruyko, S.; and Komodakis, N. 2016. Wide residual networks. *arXiv preprint arXiv:1605.07146*.

Zavrtanik, V.; Kristan, M.; and Skočaj, D. 2021a. Draem-a discriminatively trained reconstruction embedding for surface anomaly detection. In *Proceedings of the IEEE/CVF international conference on computer vision*.

Zavrtanik, V.; Kristan, M.; and Skočaj, D. 2021b. Reconstruction by inpainting for visual anomaly detection. *Pattern Recognition*.

Zhang, C.; Chang, C.-c.; and Jamshidi, M. 2020. Concrete bridge surface damage detection using a single-stage detector. *Computer-Aided Civil and Infrastructure Engineering*.

Zhang, H.; Wu, Z.; Wang, Z.; Chen, Z.; and Jiang, Y.-G. 2023a. Prototypical residual networks for anomaly detection and localization. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*.

Zhang, X.; Li, S.; Li, X.; Huang, P.; Shan, J.; and Chen, T. 2023b. Destseg: Segmentation guided denoising student-teacher for anomaly detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*.

Zhang, X.; Xu, M.; and Zhou, X. 2024. RealNet: A feature selection network with realistic synthetic anomaly for anomaly detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*.

Zou, Q.; Zhang, Z.; Li, Q.; Qi, X.; Wang, Q.; and Wang, S. 2018. Deepcrack: Learning hierarchical convolutional features for crack detection. *IEEE transactions on image processing*.

Zou, Y.; Jeong, J.; Pemula, L.; Zhang, D.; and Dabeer, O. 2022. Spot-the-difference self-supervised pre-training for anomaly detection and segmentation. In *European Conference on Computer Vision*. Springer.