

## ノイズとして異常を除去する産業用欠陥局所化

Fanbin Lu<sup>1</sup>   Xufeng Yao<sup>1</sup>   Chi-Wing Fu<sup>1</sup>   Jiaya Jia<sup>1,2</sup>  
<sup>1</sup>The Chinese University of Hong Kong   <sup>2</sup>SmartMore  
 {fblu21, xfyao, cwfu, leojia}@cse.cuhk.edu.hk

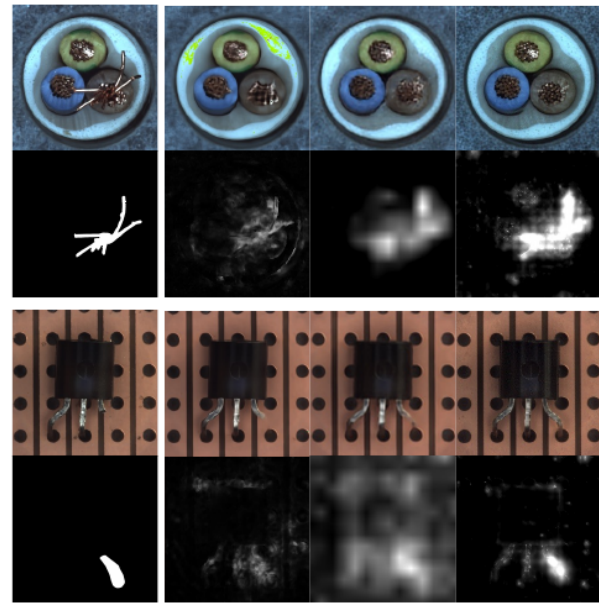
## Abstract

非監督型異常検出は、異常のない画像のみを使用してモデルを訓練し、未見の異常を検出・局所化することを目的とします。従来の再構築ベースの手法は、不正確な再構築結果に制限されていました。本研究では、生成型拡散モデルを用いた異常検出・局所化のためのノイズ除去モデルを提案します。特に、異常ピクセルを覆い隠すためのランダムノイズを導入し、中間ノイズ除去プロセスからピクセル単位の正確な異常スコアを取得します。拡散モデルのKLダイバージェンスは、伝統的なRGB空間スコアよりも優れた異常スコアとして機能することが判明しました。さらに、事前訓練された深層特徴抽出器から特徴を再構築し、特徴レベルスコアとして用いることで、局所化性能を向上させます。さらに、異常画像を正常画像に滑らかに変換する勾配ノイズ除去プロセスを提案します。当研究のノイズ除去モデルは、MVTec-ADベンチマークにおいて、正確な異常局所化と高品質な正常画像再構築において、最先端の再構築ベース異常検出手法を凌駕する性能を示しました。

## 1. Introduction

異常検出は、産業と医療において大きな応用価値を有する重要なコンピュータビジョントaskです。その重要性にもかかわらず、異常データの収集とアノテーションは極めて高額になる可能性があります。最近、無監督異常検出が注目を浴びています。少例セグメンテーション[12, 33, 32]とは異なり、訓練時に異常サンプルや ground-truth アノテーションにアクセスせずに、正常データの分布を学習することを目的としています。推論時、異常は学習した正常データの分布からの偏差に基づいて検出され、局所化されます。

古典的な再構築ベースの非監督型異常検出手法 [1, 2, 5, 20] は、正常データのみで訓練されたオートエンコーダーモデルが異常領域を再構築できないと仮定しています。しかし、このアプローチには限界があり、一部の異常は依然として再構築可能であり、これにより古典的な手法の性能が低下する原因となります。



(a) Input & G.T. (b) DRAEM (c) UniAD (d) Ours

図1：異なる手法で生成された再構築された正常画像（行1と3）と異常検出結果（行2と4）の比較。当手法は、明らかなアーティファクト（DRAEM [40]）やボケ（UniAD [36]）なしに高品質な再構築を生成し、異常をより正確に検出できます。

DRAEM [40] は、異常データを異常のない状態に再構築するために、偽の異常を生成してオートエンコーダーを訓練する方法を提案しています。しかし、実際の異常が偽の異常と大きく異なる場合、その性能は低下します。医療分野の異常検出には、デノイズングオートエンコーダー [17] が使用されています。異常スコアは、入力画像と再構築画像のピクセル空間における差分によって単純に測定されます。ノイズを含む画像からの再構築は困難であり、結果に大きなノイズを導入するため、複雑な産業用異常検出と局所化には適していません。最近、[18, 22, 36, 37]では、オートエンコーダーが恒等関数に収束するのを防ぐため、再構築モデルにトランスフォーマー構造を提案しています。



# Removing Anomalies as Noises for Industrial Defect Localization

Fanbin Lu<sup>1</sup> Xufeng Yao<sup>1</sup> Chi-Wing Fu<sup>1</sup> Jiaya Jia<sup>1,2</sup>  
<sup>1</sup>The Chinese University of Hong Kong <sup>2</sup>SmartMore  
 {fblu21, xfyao, cwfu, leojia}@cse.cuhk.edu.hk

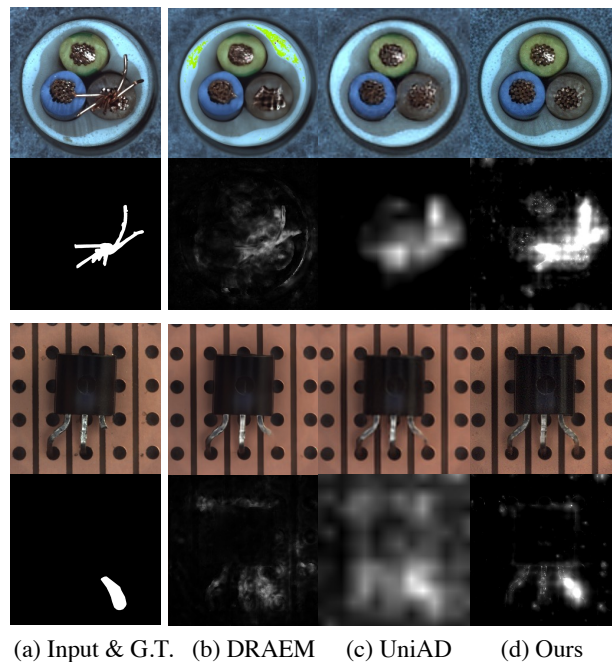
## Abstract

Unsupervised anomaly detection aims to train models with only anomaly-free images to detect and localize unseen anomalies. Previous reconstruction-based methods have been limited by inaccurate reconstruction results. This work presents a denoising model to detect and localize the anomalies with a generative diffusion model. In particular, we introduce random noise to overwhelm the anomalous pixels and obtain pixel-wise precise anomaly scores from the intermediate denoising process. We find that the KL divergence of the diffusion model serves as a better anomaly score compared with the traditional RGB space score. Furthermore, we reconstruct the features from a pre-trained deep feature extractor as our feature level score to improve localization performance. Moreover, we propose a gradient denoising process to smoothly transform an anomalous image into a normal one. Our denoising model outperforms the state-of-the-art reconstruction-based anomaly detection methods for precise anomaly localization and high-quality normal image reconstruction on the MVTec-AD benchmark.

## 1. Introduction

Anomaly detection is a critical computer vision task that has great application values in industry and medicine. Despite its importance, collecting and annotating anomalous data can be prohibitively expensive. Unsupervised anomaly detection recently garnered significant attention. Different from few-shot segmentation [12, 33, 32], it aims to learn normal data distribution without access to anomalous samples and ground-truth annotations in training. At inference, anomalies are detected and localized based on their deviation from the learned distribution of normal data.

Classical reconstruction-based unsupervised anomaly detection methods [1, 2, 5, 20] assume the autoencoder model trained with only normal data fail to reconstruct anomalous regions. However, this approach is not without limitations, as some anomalies can still be reconstructed, leading to the inferior performance of these classical methods. DRAEM [40] proposes to generate pseudo anomalies



(a) Input & G.T. (b) DRAEM (c) UniAD (d) Ours

Figure 1: Comparing reconstructed normal images (rows 1 & 3) and anomaly detection results (rows 2 & 4) produced by different methods. Our method can produce high-quality reconstructions without obvious artifacts (DRAEM [40]) and blurring (UniAD [36]) and locate anomalies more precisely.

to train an autoencoder to reconstruct the anomalous data to be anomaly-free. However, it performs poorly when the real anomalies differ significantly from the pseudo ones. Denoising autoencoders [17] are used for medical anomaly detection. The anomaly score is measured naively by the difference between the input and reconstructed images in pixel space. The reconstruction from noisy images is challenging and introduces great noise to the results, making it unsuitable for complex industrial anomaly detection and localization. Recently, methods [18, 22, 36, 37] propose transformer structures for the reconstruction model to prevent the autoencoder from collapsing into an identity func-