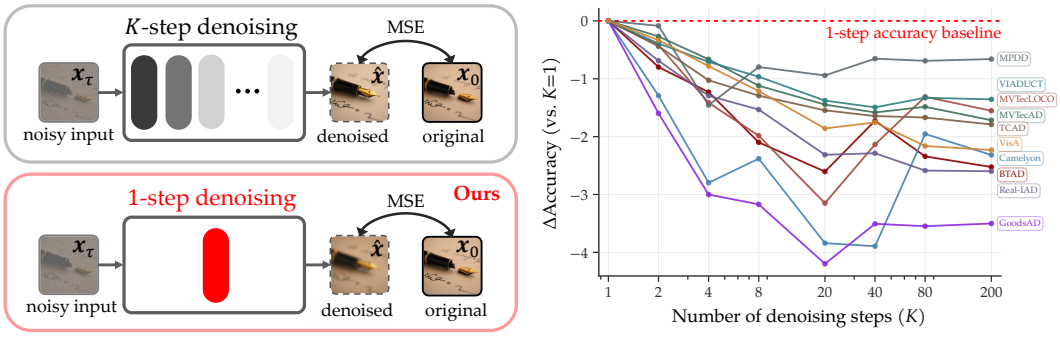
DIFFUSION MODELS ARE SECRETLY ONE-STEP ANOMALY DETECTORS

Anonymous authors
Paper under double-blind review

ABSTRACT

Diffusion-based anomaly detection (AD) commonly relies on iterative denoising to obtain high-quality reconstructions, following the intuition that better reconstruction leads to better AD. In this work, we revisit this intuition and ask whether iterative refinement is truly beneficial for AD. Through a large-scale empirical study across diverse AD datasets and diffusion designs, we uncover a striking phenomenon that we term *one-step optimality*: diffusion models trained solely on normal samples often achieve their best detection performance with just a single denoising step. To understand why, we theoretically analyze the optimal velocity field and show that iterative denoising can increasingly reconstruct anomalous information, leading to greater *anomaly leakage* and reduced separation between normal and anomaly scores. Under our analysis setting, one-step denoising minimizes such leakage and achieves optimal detection performance. This understanding naturally leads to the *One-step Diffusion anomaly Detector (ODD)*, which retains standard diffusion training but performs reconstruction with a single denoising step. Extensive experiments show that **ODD** achieves strong detection performance with substantially improved inference efficiency across diverse datasets and modalities, while also serving as a plug-and-play replacement for iterative denoising in existing diffusion-based AD methods. Code is available at: https://anonymous.4open.science/r/odd.



(a) One-step vs. multi-step denoising **(b) Detection performance vs. denoising steps**

Figure 1: **One-step optimality in diffusion-based anomaly detection. (a)** Conventional diffusion-based AD relies on multi-step denoising to progressively refine reconstructions, whereas we find that a single denoising step suffices for effective anomaly detection. **(b)** Across diverse AD benchmarks, detection performance is mostly highest with just a single denoising step.

1 INTRODUCTION

Anomaly detection (AD) plays a pivotal role in industrial defect detection (Bergmann et al., 2019), medical diagnosis (Bao et al., 2024), and video surveillance (Zhu et al., 2024). In realistic settings, anomalous samples are often scarce, making unsupervised AD particularly important because it can learn the normal data distribution and identify deviations from it as anomalies (Ruff et al., 2021; Xie et al., 2024; Liu et al., 2024). For unsupervised AD, generative models based on *denoising*, such as diffusion models (Sohl-Dickstein et al., 2015; Ho et al., 2020) and flow matching models (Albergo & Vanden-Eijnden, 2023; Lipman et al., 2023; Liu et al., 2023b), have attracted substantial interest for

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

their ability to characterize high-dimensional and complex data distributions. Recent methods have shown that diffusion models trained solely on normal samples cannot precisely reconstruct anomalous regions from noisy test samples, thereby conveniently enabling AD by measuring reconstruction errors (Li et al., 2024c; Fučka et al., 2025; Feng et al., 2025; He et al., 2024b).

Existing diffusion-based AD methods commonly rely on iterative denoising to progressively refine reconstructions (Lu et al., 2023a; Li et al., 2024c; Yao et al., 2025), as illustrated in Figure 1(a, top). This design follows the intuition that more thoroughly denoised, higher-quality reconstructions are beneficial for AD, and multi-step denoising is therefore adopted at inference. However, AD differs fundamentally from conventional generation: its goal is not to produce high-quality reconstructions, but to distinguish anomalous samples from normal ones through their reconstruction errors. This raises a fundamental question: *is iterative refinement actually beneficial for AD?*

In this paper, we systematically investigate this question across diverse AD tasks, including visual defect detection (Bergmann et al., 2019; Jezek et al., 2021; Zou et al., 2022; Wang et al., 2024), medical imaging analysis (Bao et al., 2024), and video surveillance (Liu et al., 2018; Acsintoae et al., 2022). Our large-scale empirical analysis uncovers a striking phenomenon, which we term *one-step optimality*, across diverse diffusion architectures, latent representations, diffusion formulations, noise strengths, and data modalities: diffusion models trained solely on normal samples often achieve their best AD performance with just a single denoising step. To gain more insights, we focus on the optimal velocity field (Bertrand et al., 2025) and theoretically show that iterative denoising can increasingly reconstruct anomalous information, leading to greater *anomaly leakage* and reducing the discriminability between normal and anomalous samples, whereas one-step denoising minimizes such leakage under our analysis setting. Building on these empirical and theoretical findings, we introduce **O**ne-step **D**iffusion anomaly **D**etector (**ODD**), a simple diffusion-based AD framework that translates *one-step optimality* into practice and can be readily applied to existing diffusion-based AD methods in a plug-and-play manner Figure 1(a, bottom). We highlight our main contributions:

- We uncover *one-step optimality* in diffusion-based AD through large-scale empirical analysis, demonstrating that a single denoising step often achieves the best detection performance across diverse diffusion designs and data modalities.
- We theoretically explain *one-step optimality* through optimal velocity fields, characterizing its connection to anomaly leakage and detection discriminability.
- We introduce *ODD*, which turns this finding into a simple one-step AD framework, achieving strong detection performance and substantially improved inference efficiency.

2 Preliminaries and Problem Setup

2.1 Problem Statement

Problem setting. We consider an unsupervised AD scenario, where only normal samples are available during training. Let $\mathcal{X} = \{\boldsymbol{x}^{(i)}\}_{i=1}^{N} \subset \mathbb{R}^d$ be a dataset of d-dimensional observations containing N samples in total, partitioned into a training set $\mathcal{X}_{\text{train}}$ containing only normal instances sampled from a distribution $Q_{\text{normal}}(\boldsymbol{x})$, and a test set $\mathcal{X}_{\text{test}}$ that may include both normal and anomalous samples.

2.2 Diffusion Models for Anomaly Detection

Basics of diffusion modeling. Diffusion models (Sohl-Dickstein et al., 2015; Ho et al., 2020) and flow matching (Lipman et al., 2023; Liu et al., 2023b) provide a principled and flexible way for learning high-dimensional data distributions. Despite their different formulations, both support reconstruction from corrupted observations through a learned reverse process, making them well suited to reconstruction-based AD. For theoretical clarity, we use flow matching as our primary formulation, as it provides a simple and analytically tractable velocity-field representation. We later theoretically show that our key finding extends to other diffusion formulations (see Appendix B).

Consider a data distribution $\boldsymbol{x} \sim P_0(\boldsymbol{x})$ and a source distribution $\boldsymbol{\epsilon} \sim P_1(\boldsymbol{\epsilon})$ (*e.g.*, $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I})$). Flow matching defines an interpolation between data and noise as:

$$\boldsymbol{x}_t = (1-t)\boldsymbol{x}_0 + t\boldsymbol{\epsilon}, \tag{1}$$

which yields $\boldsymbol{x}_t \sim P_0$ at $t = 0$ and $\boldsymbol{x}_t \sim P_1$ at $t = 1$. A key quantity in flow matching is *velocity*, which is defined as the time derivative of $\boldsymbol{x}_t$, that is, $\boldsymbol{v} := d\boldsymbol{x}_t/dt = \boldsymbol{\epsilon} - \boldsymbol{x}_0$. A parameterized velocity field $\boldsymbol{v}_\theta$ is learned by minimizing

$$\mathcal{L}_{\text{fm}}(\theta) := \mathbb{E}_{\boldsymbol{x}_0 \sim P_0, \boldsymbol{\epsilon} \sim P_1, t \sim U[0,1]}[||\boldsymbol{v}_\theta(\boldsymbol{x}_t, t) - (\boldsymbol{\epsilon} - \boldsymbol{x}_0)||_2^2]. \tag{2}$$

During inference, we can generate samples by solving an ordinary differential equation (ODE) with initial condition $\boldsymbol{x}_1 \sim P_1$, integrated backward in time from $t = 1$ to $t = 0$:

$$\boldsymbol{x}_0 = \boldsymbol{x}_1 + \int_1^0 \boldsymbol{v}_\theta(\boldsymbol{x}_t, t)dt. \tag{3}$$

Optimal velocities. In practice, we approximate the true data distribution Q_{normal} by the empirical distribution consisting of Dirac measures $Q_{\text{emp}} = \frac{1}{n}\sum_{i=1}^n \delta(\boldsymbol{x} - \boldsymbol{x}^{(i)})$ because we only have a finite training set $\mathcal{X}_{\text{train}}$, and (2) thereby reduces to its empirical counterpart. Minimizing this reduced empirical loss leads to the following nonparametric *optimal velocity* $\boldsymbol{v}^*(\boldsymbol{x}, t)$:

$$\boldsymbol{v}^*(\boldsymbol{x}, t) = \mathbb{E}_{\boldsymbol{x}_0 \sim P_0, \boldsymbol{\epsilon} \sim P_1}\left[\boldsymbol{\epsilon} - \boldsymbol{x}_0 \mid \boldsymbol{x}_t = \boldsymbol{x}\right] = \sum_{i=1}^n \gamma_i(\boldsymbol{x}, t)\frac{\boldsymbol{x} - \boldsymbol{x}^{(i)}}{t}, \tag{4}$$

where $\gamma_i(\boldsymbol{x}, t) := \text{softmax}(-\frac{||\boldsymbol{x}-(1-t)\boldsymbol{x}^{(i)}||_2^2}{2t^2})$ (Bertrand et al., 2025; Gu et al., 2025). While (4) provides an analytical form of the sampling process of the flow matching, it provably only produces training samples and never generalizes (Zhang et al., 2024a; Baptista et al., 2025; Bertrand et al., 2025; Gu et al., 2025). On the other hand, we observe that the optimal velocity also exhibits a similar optimality of one-step denoising for AD (Figure 2(e)). Therefore, we first analyze the optimal velocity, then extend the discussion to a realistic learned velocity (See Appendix A).

Anomaly detection with diffusion models. In reconstruction-based AD, a denoising model trained only on normal data is expected to reconstruct normal patterns more faithfully than anomalous ones, allowing anomalies to be identified through reconstruction errors. Given a test sample $\boldsymbol{x}$, we first perturb it to a pre-defined noise level τ with (1), obtaining $\boldsymbol{x}_\tau$. We then reconstruct $\boldsymbol{x}_\tau$ by numerically solving the reverse-time ODE in (3) from $t = \tau$ to $t = 0$. For example, using a K-step explicit Euler method, the reconstruction process is:

$$\boldsymbol{x}_{t_0} = \boldsymbol{x}_\tau, \qquad \boldsymbol{x}_{t_{k+1}} = \boldsymbol{x}_{t_k} - h_k \boldsymbol{v}_\theta(\boldsymbol{x}_{t_k}, t_k), \quad k = 0, \ldots, K-1, \qquad \hat{\boldsymbol{x}}_{\theta,\Pi_K}(\boldsymbol{x}_\tau) := \boldsymbol{x}_{t_K}, \tag{5}$$

where $\Pi_K := (\tau = t_0 > t_1 > \cdots > t_K = 0)$ denotes a reverse-time discretization of $[0, \tau]$, and $h_k := t_k - t_{k+1} > 0$ is the corresponding step size. The anomaly score is then defined as MSE:

$$s_{\tau,\Pi_K}(\boldsymbol{x}) := \|\boldsymbol{x} - \hat{\boldsymbol{x}}_{\theta,\Pi_K}(\boldsymbol{x}_\tau)\|_2^2. \tag{6}$$

In practice, diffusion-based AD commonly adopts multiple denoising steps to progressively refine the reconstruction. This raises the question posed in §1: *does such iterative refinement actually benefit AD?* We next systematically examine this question by varying the number of denoising steps K.

3 Findings: One-Step Optimality

Across a broad range of AD benchmarks and diffusion configurations, our large-scale empirical analysis reveals a striking phenomenon: a single denoising step often achieves the best detection performance. We term this phenomenon *one-step optimality*. We next examine the robustness and generality of this phenomenon from four complementary perspectives. Following the standard diffusion-based reconstruction AD setup (Lu et al., 2023a; Yao et al., 2025), we measure how detection accuracy [1] varies across the number of denoising steps (K), relative to the $K = 1$ setting.

Finding 1: One-step optimality broadly emerges across image AD benchmarks. We examine various widely used image AD benchmarks, including industrial defect detection (Bergmann et al., 2019), compositional defect detection (Bergmann et al., 2022a), and medical analysis (Bao et al.,

[1] In this paper, following (Sakai et al., 2026), we define Accuracy (*e.g.*, in Figure 1, Figure 2) as the average over all broadly used metrics: image- and pixel-level AUROC (Area Under the Receiver Operating Characteristic Curve), AP (Average Precision), PRO (Per-Region Overlap Curve) and F1_max (Maximum F1-score).

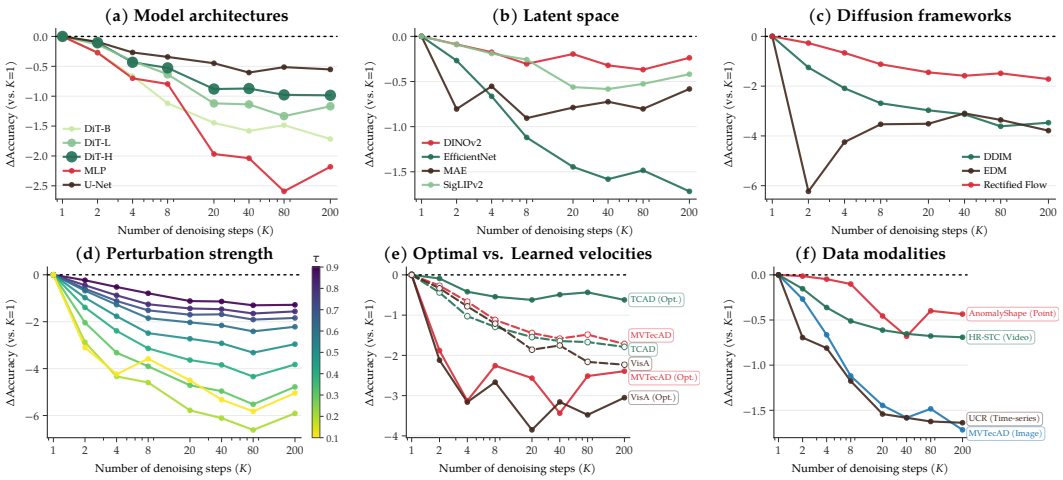


Figure 2: **Large-scale analysis reveals that one-step optimality holds across various conditions.**

2024) (see Appendix C). As seen in Figure 1(b), *one-step optimality* mostly holds regardless of diverse dataset attributes. We also provide examples on other datasets in §E.1.

Finding 2: One-step optimality persists across diverse diffusion designs. To further understand whether the *one-step optimality robustly* emerges across different designs of diffusion modeling, we systematically probe the following components in Figure 2(a)-(d): *model architectures*, *latent space*, *diffusion frameworks*, and *perturbation strength*. *One-step optimality* consistently persists across these variations, suggesting that the phenomenon is not tied to a particular diffusion design.

Finding 3: One-step optimality holds in the memorization regime. A common perception is that *one-step optimality* only holds in the generalization regime of diffusion models, as discussed in (Baptista et al., 2025; Bonnaire et al., 2025). However, in diffusion-based AD, we find that *one-step optimality* also holds in the memorization regime (Gu et al., 2025) with the optimal velocity (4). As seen in Figure 2(e), we compare reconstruction with the learned diffusion model and optimal velocity (Opt.), confirming that the optimal velocity again shows *one-step optimality* as a learned counterpart.

Finding 4: One-step optimality generalizes across data modalities. To investigate how consistently *one-step optimality* holds across modalities beyond images, we consider AD in video, point cloud, and timeseries (see Appendix C). As shown in Figure 2(f), using one representative dataset per modality, we observe that one-step denoising is optimal across various modalities, suggesting that *one-step optimality* can generalize beyond image-based AD.

Taken together, these findings establish *one-step optimality* as a broadly observed phenomenon in diffusion-based AD, rather than an artifact of a particular dataset, diffusion design, velocity field, or data modality. Having established its empirical robustness, we next investigate why one-step denoising can be optimal through the lens of optimal velocity fields (§4).

4 Understanding One-Step Optimality

In this section, we systematically show that one-step denoising is natively optimal for AD through the lens of an optimal velocity. Our key observation is that the optimality also arises from the optimal velocity, not only from the learned velocity (Figure 2(e)). Through the lens of the optimal velocity, we establish two complementary forms of *one-step optimality* by analyzing: *reconstruction error on normal samples* (§4.1) and *error gap between normal and anomalous samples* (§4.3). By combining these two optimality results, we uncover that one-step denoising is provably optimal for AD, providing a principled way towards accurate and efficient diffusion-based AD.

4.1 One-Step Optimality for Normal Samples

It is mostly prevalent to follow the paradigm of *minimizing the reconstruction error on normal samples* to reduce false positives. We first demonstrate that one-step denoising minimizes reconstruction error

with the optimal velocity if we adopt MSE as the measure of reconstruction error, which is a standard choice for previous methods (He et al., 2024b; Fučka et al., 2025).

Proposition 4.1 (One-step denoising is the MMSE estimator). *Let $\boldsymbol{x}_0 \sim Q_{normal}$ follow an arbitrary distribution Q_{normal} satisfying $\mathbb{E}[||\boldsymbol{x}_0||_2^2] < \infty$, and consider linear probability path (1) and population-optimal velocity $\boldsymbol{v}^\star(\boldsymbol{x}, t) := \mathbb{E}\left[\boldsymbol{\epsilon} - \boldsymbol{x}_0 \mid \boldsymbol{x}_t = \boldsymbol{x}\right]$ that minimizes the flow matching loss (2). Then, for any $\tau \in (0, 1]$, the one-step reverse-Euler reconstruction $\hat{\boldsymbol{x}}_{\Pi_1}(\boldsymbol{x}_\tau) := \boldsymbol{x}_\tau - \tau \boldsymbol{v}^\star(\boldsymbol{x}_{\boldsymbol{\tau}}, \tau)$ is the minimum mean square error (MMSE) estimator of $\boldsymbol{x}_0$ given $\boldsymbol{x}_\tau$:*

$$\hat{\boldsymbol{x}}_{\Pi_1} \in \operatorname*{argmin}_g \mathbb{E}[||\boldsymbol{x}_0 - g(\boldsymbol{x}_\tau)||_2^2], \tag{7}$$

where the minimum is taken over all measurable reconstruction mappings g.

Remarks. The proof is provided in Appendix §A.2. (7) indicates that the one-step denoising using the optimal velocity $\hat{\boldsymbol{x}}_{\Pi_1}$ always minimizes averaged reconstruction MSE across all normal samples, compared with all other multi-step denoisings $\hat{\boldsymbol{x}}_{\Pi_{K>1}}$. However, Proposition 4.1 does not explain the optimality for AD where the test set contains both normal and *anomalous* samples. Therefore, we next introduce the basic settings (§4.2) on data distribution and velocity model to analyze the behavior of the diffusion-based AD in a tractable way.

4.2 Analysis Setup for Anomaly Detection

Data assumptions. We assume an M-component Gaussian mixture model (GMM) for the normal data distribution[2]:

$$\boldsymbol{x}_0 \sim Q_{\text{normal}} := \sum_{j=1}^{M} w_j \mathcal{N}(\boldsymbol{\mu}_j, \lambda \boldsymbol{I}), \quad \lambda, w_j > 0, \quad \sum_{j=1}^{M} w_j = 1, \tag{8}$$

which is a standard approximation to data manifolds frequently used in recent theoretical studies (Bonnaire et al., 2025; Ye et al., 2026). To analyze the AD capability of diffusion models, we also hypothesize that an anomalous sample $\boldsymbol{x}^A$ can be represented as: $\boldsymbol{x}^A = \boldsymbol{x}^N + \boldsymbol{a}$, where $\boldsymbol{x}^N \sim Q_{\text{normal}}$ represents a counterfactual normal sample, and $\boldsymbol{a}$ vectorizes an anomaly residual in the manner of additive modeling. Indeed, various real-world anomalies (*e.g.*, industrial defects, lesions in X-rays) adopt this hypothesis, and it motivates many prior AD works in a principled mathematical setup (Datta et al., 2022; Ji et al., 2024; Cai et al., 2024; Yao et al., 2025; Dai et al., 2025).

Under the above data assumptions, the optimal velocity fields can be represented using GMM components that enable us to analyze their behavior in the tractable regime.

Proposition 4.2 (Optimal velocities under GMM-formed data distribution). *Let an M-component GMM generate normal samples $\boldsymbol{x}_0$ as in (8), and consider the linear probability path (1) which smoothly interpolates between Q_{normal} and $P_1 = \mathcal{N}(\mathbf{0}, I)$. Then, an optimal velocity (4) under the GMM-formed data distribution is:*

$$\boldsymbol{v}^\star(\boldsymbol{x}, t) = \sum_{j=1}^{M} \gamma_j(\boldsymbol{x}, t) \boldsymbol{v}^{(j)}(\boldsymbol{x}, t), \quad \boldsymbol{v}^{(j)}(\boldsymbol{x}, t) = -\boldsymbol{\mu}_j + b_t \left(\boldsymbol{x} - (1-t)\boldsymbol{\mu}_j\right), \tag{9}$$

where $\gamma_j(\boldsymbol{x}, t) = \operatorname{softmax}\left(\log w_j - \frac{\|\boldsymbol{x} - (1-t)\boldsymbol{\mu}_j\|_2^2}{2c_t}\right)$, $c_t = (1-t)^2\lambda + t^2$, $\quad b_t = \frac{t - (1-t)\lambda}{c_t}$.

The proof is provided in Appendix §A.3. Precisely, (9) is a specific instance of (4) under the GMM data distribution, and it computes a marginal velocity as a weighted sum of per-component conditional velocities, where *responsibilities* γ_j control the weight.

4.3 One-step Optimality for Anomaly Detection

Building on the setup in §2 and §4.2, let a normal sample and its anomalous counterpart share the corruption noise, and let $j^\star$ be the component generating the normal sample. We assume *posterior*

[2] In practice, most of the AD methods operate on the latent space of pre-trained networks, which can be well approximated by a GMM with very few components (Lee et al., 2018; Defard et al., 2021)

anomaly invariance (Assumption 4.3), *i.e.*, the anomaly does not alter the responsibility $\gamma_j(\boldsymbol{x}, t)$ in (9) along the reconstruction trajectory with optimal velocities (4), and *prototype fidelity* (Assumption 4.4), *i.e.*, the responsibility-weighted average of the component means recovers $\boldsymbol{\mu}_{j^\star}$. We justify both in Appendix §A.4. Under Assumption 4.3, we first show that the number of denoising steps affects the degree of anomaly leakage, *i.e.*, anomalous regions being unexpectedly reconstructed (Theorem 4.5). Under Assumption 4.4, we then relate it to the anomaly scores (Lemma 4.6) and obtain *one-step optimality* in terms of a practical AD accuracy metric, AUROC (Theorem 4.7).

Assumption 4.3 (Posterior anomaly invariance)**.** Let $\boldsymbol{x}_\tau^N = (1-\tau)\boldsymbol{x}_0^N + \tau\boldsymbol{\epsilon}$ and $\boldsymbol{x}_\tau^A = (1-\tau)\boldsymbol{x}_0^A + \tau\boldsymbol{\epsilon}$ with $\boldsymbol{x}_0^A = \boldsymbol{x}_0^N + \boldsymbol{a}$, and let further $\{\boldsymbol{x}_{t_k}^N\}_k$ and $\{\boldsymbol{x}_{t_k}^A\}_k$ be their K-step Euler reconstruction trajectories. Then $\gamma_j(\boldsymbol{x}_{t_k}^A, t_k) = \gamma_j(\boldsymbol{x}_{t_k}^N, t_k)$ for all j and $k = 0, \ldots, K-1$.

Assumption 4.4 (Prototype fidelity)**.** Under the notation of Assumption 4.3, $\bar{\boldsymbol{m}}_{\Pi_K}(\boldsymbol{x}_\tau^N) = \bar{\boldsymbol{m}}_{\Pi_K}(\boldsymbol{x}_\tau^A) = \boldsymbol{\mu}_{j^\star}$ holds almost surely, where $\bar{\boldsymbol{m}}_{\Pi_K}$ is the prototype defined as:

$$\bar{\boldsymbol{m}}_{\Pi_K}(\boldsymbol{x}_\tau) := \sum_{k=0}^{K-1} \beta_k \, \boldsymbol{m}(\boldsymbol{x}_{t_k}, t_k), \qquad \boldsymbol{m}(\boldsymbol{x}, t) := \sum_{j=1}^{M} \gamma_j(\boldsymbol{x}, t)\boldsymbol{\mu}_j, \tag{10}$$

where $\beta_k > 0$ is a step-dependent scalar satisfying $\sum_k \beta_k = 1$ (see Lemma A.1 for a definition).

Theorem 4.5 (Anomaly leakage from optimal velocities)**.** *Let $\boldsymbol{x}_\tau^N$ and $\boldsymbol{x}_\tau^A$ be as in Assumption 4.3, and consider the reconstruction operator $\hat{\boldsymbol{x}}_{\Pi_K}$ derived by the optimal velocity (4). Under Assumption 4.3, it holds:*

$$\hat{\boldsymbol{x}}_{\Pi_K}(\boldsymbol{x}_\tau^A) - \hat{\boldsymbol{x}}_{\Pi_K}(\boldsymbol{x}_\tau^N) = \ell_{\Pi_K}\boldsymbol{a}, \tag{11}$$

where $\ell_{\Pi_K} := (1-\tau)\prod_{k=0}^{K-1}(1 - h_k b_{t_k})$.

Remarks. The proof is provided in Appendix §A.5. (11) reveals that if the responsibilities remain unaltered by the anomaly along the temporal reconstruction trajectory, the difference between normal and anomalous reconstructions can be represented as a scaled anomaly signal where the number of denoising steps K **only** affects the scalar coefficient $\ell_{\Pi_K} \in [0, 1)$. As the anomalous signal is expected to be suppressed through the reconstruction process $\hat{\boldsymbol{x}}_{\Pi_K}(\boldsymbol{x}_\tau^A)$ for better discriminability, minimizing ℓ_{Π_K} directly contributes to improved AD accuracy.

Next, we connect ℓ_{Π_k} to a noncentrality parameter of the distribution of anomaly scores.

Lemma 4.6 (Reconstruction error follows χ^2)**.** *Suppose that Assumption 4.4 holds for a normal sample $\boldsymbol{x}_0^N$ and its anomalous counterpart $\boldsymbol{x}_0^A = \boldsymbol{x}_0^N + \boldsymbol{a}$, where $\boldsymbol{a} \neq \boldsymbol{0}$ is a fixed vector. Then, anomaly scores $S_{\Pi_K}^N$ and $S_{\Pi_K}^A$ follow chi-square distributions χ_d^2 and $\chi_d'^2$, respectively:*

$$\frac{S_{\Pi_K}^N}{\sigma_{\Pi_K}^2} \sim \chi_d^2(0), \qquad \frac{S_{\Pi_K}^A}{\sigma_{\Pi_K}^2} \sim \chi_d'^2(\Lambda_{\Pi_K}), \tag{12}$$

where $\sigma_{\Pi_K}^2 := \lambda(1-\ell_{\Pi_K})^2 + {}^{\tau^2\ell_{\Pi_K}^2}/_{(1-\tau)^2}$ is the variance of each entry, and $\Lambda_{\Pi_K} := \|\boldsymbol{a}\|_2^2\big[\lambda + \tau^2\ell_{\Pi_K}^2/\{(1-\tau)(1-\ell_{\Pi_K})\}^2\big]^{-1}$ is the noncentrality parameter of the $\chi_d'^2$.

Theorem 4.7 (One-step optimality)**.** *Under the setup of Lemma 4.6, one-step reconstruction strictly outperforms any multi-step reconstruction:*

$$\mathrm{AUROC}_{\Pi_1} > \mathrm{AUROC}_{\Pi_K} \quad \forall K > 1. \tag{13}$$

Proof sketch (See Appendix A.6 for the complete proof). As both scores share the scale $\sigma_{\Pi_K}^2$, the AUROC is strictly increasing in Λ_{Π_K}, and Λ_{Π_K} is strictly decreasing in ℓ_{Π_K}, and $\ell_{\Pi_1} < \ell_{\Pi_K}$. □

Remarks. Under Assumption 4.4, Theorem 4.7 attributes the advantage of one-step reconstruction to its minimal anomaly leakage ℓ_{Π_1}. While Proposition 4.1 considers only normal samples, this result holds even when anomalous samples are taken into account.

5 ODD: PUTTING ONE-STEP OPTIMALITY INTO PRACTICE

Building on the empirical (§3) and theoretical (§4) evidence of **one-step optimality**, we introduce **O**ne-step **D**iffusion anomaly **D**etector (**ODD**), which directly translates this finding into a practical AD framework. The key idea is simple: just retain the standard diffusion training procedure, but replace iterative reconstruction with a single denoising step:

$$\hat{\boldsymbol{x}}(\boldsymbol{x}_\tau) := \boldsymbol{x}_\tau - \tau \boldsymbol{v}_\theta(\boldsymbol{x}_\tau, \tau). \tag{14}$$

Concretely, **ODD** is nothing more than substituting (3) with (14) for reconstruction, with all other AD-related components retained, including the diffusion implementation and scoring policy. This simple design not only realizes the detection benefit of *one-step optimality*, but also substantially reduces inference cost by requiring only a single NFE. Moreover, it yields several desirable properties: plug-and-play, architecture robustness, and modality generality, which we will demonstrate in §6.

6 EXPERIMENTS

6.1 EXPERIMENTAL SETTINGS

Datasets. To measure detection accuracy on various real-world AD tasks, we utilize broadly used image-based AD datasets: MVTecAD (Bergmann et al., 2019), VisA (Zou et al., 2022), MPDD (Jezek et al., 2021), and RealIAD (Wang et al., 2024). We also evaluate our proposed framework on other modalities: HR-STC (Liu et al., 2018), HR-UBN (Acsintoae et al., 2022) for video AD; Real3D-AD (Liu et al., 2023a), AnomalyShape (Li et al., 2024b) for point cloud AD; SMD (Su et al., 2019), UCR (Wu & Keogh, 2022) for time-series AD (See Appendix C for dataset details).

Evaluation setups. For all experimental results, we follow the unsupervised AD setup, where only normal samples are available during training. In image AD, we report widely adopted image- and pixel-level metrics (AUROC, AP, F1_max, AU-PRO) and their average as mAD. Consistent with prior works (Zhou et al., 2025; Cheng et al., 2025; Obata et al., 2026), we report frame-wise AUROC for video, object-wise AUROC for point cloud, and VUS-AUROC (Boniol et al., 2025) for time-series. We measure frames-per-second (FPS) on an RTX4090 as in InvAD (Sakai et al., 2026).

Implementation details. We adopt the simple diffusion transformer (DiT) (Peebles & Xie, 2023) for the diffusion architecture across all experiments. To extract semantic features, we employ pre-trained EfficientNet-B4 (Tan & Le, 2019) as in previous methods (Sakai et al., 2026; Feng et al., 2025; Lu et al., 2023b), resulting in $\boldsymbol{x}_0 \in \mathbb{R}^{272\times32\times32}$. DiT is conditioned on class labels and trained for 300 epochs using the Adam (Kingma & Ba, 2015) optimizer with the standard FM training objective (2). In Appendix C, we provide a more detailed explanation of the experimental setting.

6.2 RESULTS

ODD achieves strong detection performance with one-step reconstruction. We compare ODD with prior diffusion-based multi-class AD approaches, including DiAD (He et al., 2024b), VPOT (Li et al., 2024c), OmiAD (Feng et al., 2025), DiffusionAD (Zhang et al., 2025), InvAD (Sakai et al., 2026), and ReFP-AD (Lendering et al., 2026). As shown in Table 1, **ODD** generally achieves state-of-the-art or competitive detection performance across the four image AD benchmarks, demonstrating that the empirical *one-step optimality* translates effectively into a practical AD framework. Moreover, requiring only a single NFE naturally brings substantial inference acceleration (*e.g.*, FPS: **406.0** *vs.* 159.5 on RealIAD), without relying on ad-hoc schemes (Sakai et al., 2026; Feng et al., 2025).

ODD generalizes across modalities. In Figure 2(f), we have demonstrated that *one-step optimality* for AD holds across modalities and is not restricted to images. To systematically investigate the generality of **ODD**, we evaluate on AD benchmarks in other modalities, including video (Liu et al., 2018; Acsintoae et al., 2022), point cloud (Liu et al., 2023a; Li et al., 2024b), and time-series (Wu & Keogh, 2022; Su et al., 2019). Here, we compare with representative methods on each modality: TrajRec (Stergiou et al., 2024), MoCoDAD (Flaborea et al., 2023), and DCMD (Wang et al., 2025) for video; IMRNet (Li et al., 2024b), R3DAD (Zhou et al., 2025), and MC3D-AD (Cheng et al., 2025) for point cloud; IMDiffusion (Chen et al., 2023), D^{3}R (Wang et al., 2023), and AnomalyFilter (Obata et al., 2026) for time-series. The results are presented in Table 2.

Table 1: **Quantitative results** on different AD datasets under the multi-class setting. For each row, mAD averages the seven image- and pixel-level metrics (*i.e.*, the 4-th to the 10-th columns).

Dataset	Method	NFE	Image-level			Pixel-level				mAD	FPS
		[↓]	AUROC	AP	F1_max	AUROC	AP	F1_max	AU-PRO		
MPDD	DiAD *[AAAI'24]*	10	74.6	82.1	82.5	93.0	15.9	21.2	78.4	64.0	0.1
	VPOT *[ICML'24]*	1000	77.4	80.0	83.6	95.9	16.4	22.5	85.3	65.9	3.1
	OmiAD *[ICML'25]*	**1**	93.1	92.9	90.6	98.5	36.1	40.7	**93.9**	78.0	49.8
	InvAD *[CVPR'26]*	3	96.5	96.5	94.4	98.3	40.1	43.9	91.2	80.1	120.0
	ReFP-AD *[ECCV'26]*	40	74.6	76.3	79.3	97.5	28.7	31.2	90.2	68.3	159.3
	ODD (DiT-B)	**1**	92.8	93.1	90.1	98.2	37.6	41.9	92.8	78.1	**325.5**
	ODD (DiT-G)	**1**	**97.9**	**98.4**	**97.0**	**98.8**	**45.3**	**48.2**	93.4	**82.7**	168.2
RealIAD	DiAD *[AAAI'24]*	10	75.6	66.4	69.9	88.0	2.9	7.1	58.1	52.6	0.1
	VPOT *[ICML'24]*	1000	64.3	61.1	63.1	93.7	11.1	17.9	75.1	55.2	2.6
	OmiAD *[ICML'25]*	**1**	**90.3**	**89.3**	**83.3**	99.0	**34.4**	40.4	93.1	**75.7**	52.1
	InvAD *[CVPR'26]*	3	83.8	77.1	75.6	97.1	22.6	31.6	87.7	67.9	90.5
	ReFP-AD *[ECCV'26]*	40	72.6	68.9	68.1	95.1	12.0	20.0	78.3	59.3	159.5
	ODD (DiT-B)	**1**	87.3	82.6	78.1	99.0	33.1	40.0	93.5	73.4	**406.0**
	ODD (DiT-G)	**1**	88.1	83.3	78.8	**99.1**	33.9	**40.6**	**93.8**	73.9	163.6
VisA	DiAD *[AAAI'24]*	10	86.8	88.3	85.1	96.0	26.1	33.0	75.2	70.1	0.1
	VPOT *[ICML'24]*	1000	68.2	75.2	75.1	92.7	13.8	19.9	66.1	58.7	3.2
	DiffusionAD *[TPAMI'25]*	**1**	92.5	94.4	88.8	97.6	37.6	42.7	92.1	78.1	32.3
	OmiAD *[ICML'25]*	**1**	95.4	96.6	92.0	98.8	39.3	43.6	89.4	79.3	35.3
	InvAD *[CVPR'26]*	3	**96.9**	97.2	93.7	**99.1**	39.2	43.1	**92.9**	80.3	74.1
	ReFP-AD *[ECCV'26]*	40	88.9	90.7	86.5	98.4	36.3	41.1	88.7	75.8	159.3
	ODD (DiT-B)	**1**	96.4	97.0	93.0	98.8	**40.9**	**45.5**	91.4	80.4	**322.7**
	ODD (DiT-G)	**1**	**96.9**	**97.7**	**94.1**	98.9	40.7	44.1	91.8	**80.6**	167.5
MVTecAD	DiAD *[AAAI'24]*	10	97.2	99.0	96.5	96.8	**52.6**	55.5	90.7	84.0	0.1
	VPOT *[ICML'24]*	1000	82.8	92.3	89.1	89.6	30.6	35.6	76.6	71.0	3.2
	DiffusionAD *[TPAMI'25]*	**1**	92.4	97.0	93.7	94.4	57.4	56.0	86.5	82.5	32.3
	OmiAD *[ICML'25]*	**1**	98.2	99.3	98.1	97.5	51.7	**56.2**	93.4	84.9	39.4
	InvAD *[CVPR'26]*	3	**99.0**	**99.6**	98.5	97.5	46.5	52.3	92.7	83.7	88.1
	ReFP-AD *[ECCV'26]*	40	90.8	95.7	92.9	95.4	42.0	46.2	88.0	78.7	160.7
	ODD (DiT-B)	**1**	98.4	99.4	97.6	97.4	48.9	53.4	93.0	84.0	**324.0**
	ODD (DiT-G)	**1**	98.8	99.5	**98.8**	**97.8**	51.2	55.1	**93.9**	**85.0**	164.7

Table 2: **Quantitative comparison across different data modalities.**

Video				Point Cloud				Time-series			
Method	HR-STC	HR-UBN	FPS	Method	Real3D-AD	AnomalyShape	FPS	Method	SMD	UCR	FPS
TrajRec *[WACV'24]*	77.9	68.2	**248.6**	IMRNet *[CVPR'24]*	72.5	66.1	1.1	IMDiffusion *[PVLDB'23]*	85.0	63.0	7.13
MoCoDAD *[ICCV'23]*	77.6	68.4	0.4	R3DAD *[ECCV'24]*	73.4	74.9	**15.9**	D^3R *[NeurIPS'23]*	82.8	58.0	192.7
DCMD *[AAAI'25]*	78.6	69.0	0.3	MC3D-AD *[IJCAI'25]*	78.2	84.2	12.2	AnomalyFilter *[PAKDD'26]*	89.1	72.6	7.5
ODD (Ours)	**79.1**	**69.7**	67.1	**ODD (Ours)**	**83.8**	**90.5**	15.0	**ODD (Ours)**	**89.4**	**79.3**	**484.5**

These results further confirm that **ODD** yields SoTA accuracy while maintaining competitive inference speed.

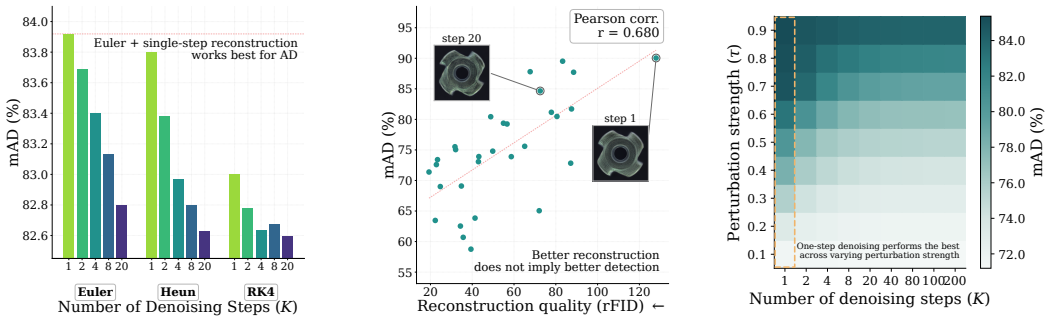
Table 3: **Plug-and-play evaluation.**

Method	mAD	FPS	NFE
VPOT *[ICML'24]*	71.0	3.2	1000
VPOT+**ODD**	**81.4**	**648.8**	**1**
InvAD *[CVPR'26]*	**83.7**	88.1	3
InvAD+**ODD**	**83.7**	**165.4**	**1**
ReFP-AD *[ECCV'26]*	78.7	160.7	40
ReFP-AD+**ODD**	**80.2**	**401.2**	**1**

ODD serves as a convenient plug-and-play module. Thanks to the simplicity of our design, **ODD** can be applied to existing AD methods in a plug-and-play fashion. To evaluate the effectiveness, we incorporate **ODD** into three representative diffusion-based AD methods: VPOT (Li et al., 2024c), InvAD (Sakai et al., 2026), and ReFP-AD (Lendering et al., 2026). As shown in Table 3, **ODD** successfully serves as a plug-and-play module to harvest improved (or maintained) accuracy and remarkable inference acceleration (*e.g.*, FPS: **165.4** *vs.* 88.1 for InvAD).

One-step optimality consistently holds across different ODE solvers. To investigate the robustness of our *one-step optimality* to solver types, we compare mAD on MVTecAD using different ODE solvers. As shown in Figure 3(a), *one-step optimality* holds across different ODE solvers. Also, we notice that the simple Euler method performs best for AD compared with other higher-order solvers, which motivates us to adopt the "Euler + one-step reconstruction" configuration in our experiments.

Better reconstruction quality does not necessarily imply better detection. To revisit the intuition motivating our study, we examine the relationship between reconstruction quality and detection performance on MVTecAD. In Figure 3(b), we plot the relationship between generation quality, measured by reconstruction Fréchet Inception Distance (rFID), and AD accuracy (mAD) across varying denoising steps. As expected, lower rFID (better reconstruction quality) does not correspond



(a) Solvers. **(b) Reconstruction quality.** **(c) Accuracy landscape.**

Figure 3: **Component analysis** on (a) Solvers, (b) Reconstruction quality, and (c) Accuracy landscape.

to higher detection accuracy; instead, rFID and mAD exhibit a positive correlation ($r = 0.680$). This clearly shows that iterative reconstruction refinement is not necessarily beneficial for AD.

One-step denoising performs the best across varying perturbation strengths. To jointly study the effects of the number of denoising steps (K) and perturbation strength (τ) on AD accuracy, we construct different (K, τ) pairs and report the corresponding mAD on MVTecAD using the pre-trained ODD (DiT-G). As shown in Figure 3(c), one-step denoising consistently achieves the best performance across the entire range of perturbation strengths, while favoring reconstruction from noisier inputs (*i.e.*, mAD: 85.0 $[\tau = 0.9]$ *vs.* 83.6 $[\tau = 0.5]$). Notably, this preference for larger τ aligns with our theoretical explanation in §A.4. We further observe that the best-performing perturbation strength remains consistently around $\tau = 0.9$ across multiple datasets, as shown in §D.2.

7 Related Work

Diffusion modeling in unsupervised anomaly detection. Given normal training samples, diffusion models (Ho et al., 2020; Lipman et al., 2023) learn a continuous mapping between a predefined noise distribution and the underlying data distribution. Such a mapping has been proven to be effective for AD on various modalities (Liu et al., 2025). Recent works (Li et al., 2024c; He et al., 2024b; Zhang et al., 2025; Feng et al., 2025; Yao et al., 2025; Lu et al., 2023a) have further exploited the potential of diffusion models within the reconstruction paradigm. Most of them propose learning an auxiliary module (He et al., 2024b; Li et al., 2024c; Zhang et al., 2025), optimizing the noise strength (Wyatt et al., 2022; Lu et al., 2023a; Fučka et al., 2025), or adopting masked diffusion modeling (Feng et al., 2025). Despite their remarkable anomaly localization capability, these methods lag largely behind non-diffusion counterparts in inference speed due to the costly *iterative denoising* (Sakai et al., 2026).

Reducing iterative denoising in diffusion-based AD. In practice, diffusion-based AD methods (He et al., 2024b; Li et al., 2024c) commonly employ multiple denoising steps to obtain clean reconstructions. Existing efforts to reduce this computational cost largely regard fewer denoising steps as an efficiency objective and introduce additional mechanisms to preserve detection performance. DiffusionAD (Zhang et al., 2025) enables denoising in a single step by relying on additional reconstruction refinement and auxiliary discriminative learning with pseudo anomalies to compensate for the resulting degradation in accuracy. Recently, Feng et al. (2025) replace iterative denoising with a one-step alternative through adversarial distillation (Zhou et al., 2024), while other works (Sakai et al., 2026; Lendering et al., 2026) resort to inversion (Zhu et al., 2016) for efficient anomaly scoring. In stark contrast, we systematically investigate the role of denoising steps themselves and uncover *one-step optimality*: one-step denoising can naturally achieve strong detection performance without modifying the original diffusion training procedure or resorting to auxiliary techniques.

8 Conclusion

In this paper, we revisit whether iterative denoising is truly beneficial for diffusion-based AD and uncover *one-step optimality*. Our theoretical analysis explains this phenomenon through anomaly leakage, and **ODD** translates this finding into a simple one-step framework with strong detection performance and substantially improved inference efficiency.

AI USE STATEMENT

In this work, we used generative AI tools to formulate mathematical claims and provide feedback on research methodology or experiments. We did not use generative AI tools to develop theoretical models or conceptual frameworks, support qualitative or thematic data analysis, or implement methods. We have reviewed all AI-assisted work. We checked LLM-generated research ideas for potential plagiarism through a manual literature survey. We take responsibility for the final content of this work, including text, claims, or artifacts produced with the aid of generative AI.

REPRODUCIBILITY STATEMENT

To support the reproducibility of our results, we provide a link to an anonymized source code: https://anonymous.4open.science/r/odd. Experimental settings, including hyperparameters, model selection criteria, evaluation metrics, and baseline configurations, are specified in Appendix C. In this paper, we use publicly available datasets, so readers can easily test our claim. Software dependencies, hardware specifications, and computational requirements are also provided in Appendix C.

REFERENCES

Andra Acsintoae, Andrei Florescu, Mariana-Iuliana Georgescu, Tudor Mare, Paul Sumedrea, Radu Tudor Ionescu, Fahad Shahbaz Khan, and Mubarak Shah. UBnormal: New benchmark for supervised open-set video anomaly detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 20143–20153, 2022.

Michael Samuel Albergo and Eric Vanden-Eijnden. Building normalizing flows with stochastic interpolants. In *International Conference on Learning Representations (ICLR)*, 2023.

Jinan Bao, Hanshi Sun, Hanqiu Deng, Yinsheng He, Zhaoxiang Zhang, and Xingyu Li. BMAD: Benchmarks for medical anomaly detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 4042–4053, 2024. doi: 10.1109/CVPRW63382.2024.00408.

Ricardo Baptista, Agnimitra Dasgupta, Nikola B. Kovachki, Assad Oberai, and Andrew M. Stuart. Memorization and regularization in generative diffusion models. *arXiv preprint arXiv:2501.15785*, 2025.

Paul Bergmann, Michael Fauser, David Sattlegger, and Carsten Steger. MVTec AD — a comprehensive real-world dataset for unsupervised anomaly detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 9584–9592, 2019. doi: 10.1109/CVPR.2019.00982.

Paul Bergmann, Kilian Batzner, Michael Fauser, David Sattlegger, and Carsten Steger. Beyond dents and scratches: Logical constraints in unsupervised anomaly detection and localization. *International Journal of Computer Vision*, 130:947–969, 2022a. doi: 10.1007/s11263-022-01578-9.

Paul Bergmann, Xin Jin, David Sattlegger, and Carsten Steger. The MVTec 3D-AD dataset for unsupervised 3D anomaly detection and localization. In *Proceedings of the 17th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications – Volume 5: VISAPP*, pp. 202–213, 2022b. doi: 10.5220/0010865000003124.

Quentin Bertrand, Anne Gagneux, Mathurin Massias, and Rémi Emonet. On the closed-form of flow matching: Generalization does not arise from target stochasticity. In *Advances in Neural Information Processing Systems*, volume 38, 2025.

Paul Boniol, Ashwin K. Krishna, Marine Bruel, Qinghua Liu, Mingyi Huang, Themis Palpanas, Ruey S. Tsay, Aaron Elmore, Michael J. Franklin, and John Paparrizos. VUS: effective and efficient accuracy measures for time-series anomaly detection. *The VLDB Journal*, 34(3):32, 2025. doi: 10.1007/s00778-025-00907-x.

Tony Bonnaire, Raphaël Urfin, Giulio Biroli, and Marc Mezard. Why diffusion models don't memorize: The role of implicit dynamical regularization in training. In *Advances in Neural Information Processing Systems*, volume 38, 2025.

Yu Cai, Hao Chen, and Kwang-Ting Cheng. Rethinking autoencoders for medical anomaly detection from a theoretical perspective. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2024*, pp. 544–554, 2024.

Yuhang Chen, Chaoyun Zhang, Minghua Ma, Yudong Liu, Ruomeng Ding, Bowen Li, Shilin He, Saravan Rajmohan, Qingwei Lin, and Dongmei Zhang. ImDiffusion: Imputed diffusion models for multivariate time series anomaly detection. *Proceedings of the VLDB Endowment*, 17(3):359–372, 2023. doi: 10.14778/3632093.3632101.

Jiayi Cheng, Can Gao, Jie Zhou, Jiajun Wen, Tao Dai, and Jinbao Wang. MC3D-AD: A unified geometry-aware reconstruction model for multi-category 3D anomaly detection. In *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*, pp. 837–845, 2025. doi: 10.24963/ijcai.2025/94. Main Track.

Wei Dai, Kai Hwang, and Jicong Fan. Unsupervised anomaly detection for tabular data using deep noise evaluation. *Proceedings of the AAAI Conference on Artificial Intelligence*, 39(11): 11553–11562, 2025. doi: 10.1609/aaai.v39i11.33257.

Debanjan Datta, Feng Chen, and Naren Ramakrishnan. Framing algorithmic recourse for anomaly detection. In *Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)*, pp. 283–293, 2022. doi: 10.1145/3534678.3539344.

Thomas Defard, Aleksandr Setkov, Angelique Loesch, and Romaric Audigier. PaDiM: A patch distribution modeling framework for anomaly detection and localization. In *Pattern Recognition. ICPR International Workshops and Challenges*, pp. 475–489, 2021.

Hanqiu Deng and Xingyu Li. Anomaly detection via reverse distillation from one-class embedding. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 9727–9736, 2022. doi: 10.1109/CVPR52688.2022.00951.

Prafulla Dhariwal and Alexander Quinn Nichol. Diffusion models beat GANs on image synthesis. In *Advances in Neural Information Processing Systems*, volume 34, 2021.

Yaoxuan Feng, Wenchao Chen, Yuxin Li, Bo Chen, Yubiao Wang, Zixuan Zhao, Hongwei Liu, and Mingyuan Zhou. OmiAD: One-step adaptive masked diffusion model for multi-class anomaly detection via adversarial distillation. In *Proceedings of the International Conference on Machine Learning (ICML)*, volume 267 of *Proceedings of Machine Learning Research*, pp. 16604–16633, 2025.

Alessandro Flaborea, Luca Collorone, Guido Maria D'Amely Di Melendugno, Stefano D'Arrigo, Bardh Prenkaj, and Fabio Galasso. Multimodal motion conditioned diffusion model for skeleton-based video anomaly detection. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 10284–10295, 2023. doi: 10.1109/ICCV51070.2023.00947.

Matic Fučka, Vitjan Zavrtanik, and Danijel Skočaj. TransFusion – a transparency-based diffusion model for anomaly detection. In *Computer Vision – ECCV 2024*, pp. 91–108, 2025.

Xiangming Gu, Chao Du, Tianyu Pang, Chongxuan Li, Min Lin, and Ye Wang. On memorization in diffusion models. *Transactions on Machine Learning Research*, 2025.

Haoyang He, Yuhu Bai, Jiangning Zhang, Qingdong He, Hongxu Chen, Zhenye Gan, Chengjie Wang, Xiangtai Li, Guanzhong Tian, and Lei Xie. MambaAD: Exploring state space models for multi-class unsupervised anomaly detection. In *Advances in Neural Information Processing Systems*, volume 37, 2024a.

Haoyang He, Jiangning Zhang, Hongxu Chen, Xuhai Chen, Zhishan Li, Xu Chen, Yabiao Wang, Chengjie Wang, and Lei Xie. A diffusion-based framework for multi-class anomaly detection. *Proceedings of the AAAI Conference on Artificial Intelligence*, 38(8):8472–8480, 2024b. doi: 10.1609/aaai.v38i8.28690.

Lars Heckler-Kram, Jan-Hendrik Neudeck, Ulla Scheler, Rebecca König, and Carsten Steger. The MVTec AD 2 dataset: Advanced scenarios for unsupervised anomaly detection. *International Journal of Computer Vision*, 134:175, 2026. doi: 10.1007/s11263-026-02743-0.

Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. In *Advances in Neural Information Processing Systems*, volume 33, 2020.

Stepan Jezek, Martin Jonak, Radim Burget, Pavel Dvorak, and Milos Skotak. Deep learning-based defect detection of metal parts: evaluating current methods in complex conditions. In *Proceedings of the International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)*, pp. 66–71, 2021. doi: 10.1109/ICUMT54235.2021.9631567.

Xiayan Ji, Anton Xue, Eric Wong, Oleg Sokolsky, and Insup Lee. AR-pro: Counterfactual explanations for anomaly repair with formal properties. In *Advances in Neural Information Processing Systems*, volume 37, 2024.

Tero Karras, Miika Aittala, Timo Aila, and Samuli Laine. Elucidating the design space of diffusion-based generative models. In Alice H. Oh, Alekh Agarwal, Danielle Belgrave, and Kyunghyun Cho (eds.), *Advances in Neural Information Processing Systems*, 2022.

Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In *International Conference on Learning Representations (ICLR)*, 2015.

Kimin Lee, Kibok Lee, Honglak Lee, and Jinwoo Shin. A simple unified framework for detecting out-of-distribution samples and adversarial attacks. In *Advances in Neural Information Processing Systems*, volume 31, pp. 7167–7177, 2018.

Jan Lehr, Jan Philipps, Alik Sargsyan, Martin Pape, and Jörg Krüger. AD3: Introducing a score for anomaly detection dataset difficulty assessment using VIADUCT dataset. In *Computer Vision – ECCV 2024*, pp. 449–464, 2025.

Camile Lendering, Erkut Akdag, Joaquín Figueira, and Egor Bondarev. ReFP-AD: Rectified flow preconditioning for energy-based anomaly detection. In *Computer Vision – ECCV 2026*, 2026.

Tianhong Li, Yonglong Tian, He Li, Mingyang Deng, and Kaiming He. Autoregressive image generation without vector quantization. In *Advances in Neural Information Processing Systems*, volume 37, 2024a.

Wenqiao Li, Xiaohao Xu, Yao Gu, Bozhong Zheng, Shenghua Gao, and Yingna Wu. Towards scalable 3D anomaly detection and localization: A benchmark via 3D anomaly synthesis and a self-supervised learning network. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 22207–22216, 2024b.

Yuxin Li, Yaoxuan Feng, Bo Chen, Wenchao Chen, Yubiao Wang, Xinyue Hu, Baolin Sun, Chunhui Qu, and Mingyuan Zhou. Vague prototype-oriented diffusion model for multi-class anomaly detection. In *Proceedings of the International Conference on Machine Learning (ICML)*, 2024c.

Yaron Lipman, Ricky T. Q. Chen, Heli Ben-Hamu, Maximilian Nickel, and Matthew Le. Flow matching for generative modeling. In *International Conference on Learning Representations (ICLR)*, 2023.

Jiaqi Liu, Guoyang Xie, Ruitao Chen, Xinpeng Li, Jinbao Wang, Yong Liu, Chengjie Wang, and Feng Zheng. Real3D-AD: A dataset of point cloud anomaly detection. In *Advances in Neural Information Processing Systems*, volume 36, 2023a. doi: 10.52202/075280-1324.

Jiaqi Liu, Guoyang Xie, Jinbao Wang, Shangnian Li, Chengjie Wang, Feng Zheng, and Yaochu Jin. Deep industrial image anomaly detection: A survey. *Machine Intelligence Research*, 21(1): 104–135, 2024.

Jing Liu, Zhenchao Ma, Zepu Wang, Chenxuanyin Zou, Jiayang Ren, Zehua Wang, Liang Song, Bo Hu, Yang Liu, and Victor C. M. Leung. A survey on diffusion models for anomaly detection. *arXiv preprint arXiv:2501.11430*, 2025.

W. Liu, W. Luo, D. Lian, and S. Gao. Future frame prediction for anomaly detection – a new baseline. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018.

Xingchao Liu, Chengyue Gong, and Qiang Liu. Flow straight and fast: Learning to generate and transfer data with rectified flow. In *International Conference on Learning Representations (ICLR)*, 2023b.

Zhikang Liu, Yiming Zhou, Yuansheng Xu, and Zilei Wang. SimpleNet: A simple network for image anomaly detection and localization. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 20402–20411, 2023c.

Ilya Loshchilov and Frank Hutter. Decoupled weight decay regularization. In *International Conference on Learning Representations (ICLR)*, 2019.

Fanbin Lu, Xufeng Yao, Chi-Wing Fu, and Jiaya Jia. Removing anomalies as noises for industrial defect localization. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 16120–16129, 2023a. doi: 10.1109/ICCV51070.2023.01481.

Ruiying Lu, YuJie Wu, Long Tian, Dongsheng Wang, Bo Chen, Xiyang Liu, and Ruimin Hu. Hierarchical vector quantized transformer for multi-class unsupervised anomaly detection. In *Advances in Neural Information Processing Systems*, volume 36, pp. 8487–8500, 2023b.

Pankaj Mishra, Riccardo Verk, Daniele Fornasier, Claudio Piciarelli, and Gian Luca Foresti. VT-ADL: A vision transformer network for image anomaly detection and localization. In *Proceedings of the IEEE International Symposium on Industrial Electronics (ISIE)*, pp. 1–6, 2021. doi: 10.1109/ISIE45552.2021.9576231.

Romero Morais, Vuong Le, Truyen Tran, Budhaditya Saha, Moussa Mansour, and Svetha Venkatesh. Learning regularity in skeleton trajectories for anomaly detection in videos. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 11988–11996, 2019. doi: 10.1109/CVPR.2019.01227.

Kohei Obata, Zheng Chen, Yasuko Matsubara, Lingwei Zhu, and Yasushi Sakurai. AnomalyFilter: Selective denoising diffusion model for time series anomaly detection. In *Advances in Knowledge Discovery and Data Mining – PAKDD 2026*, pp. 429–441, 2026. doi: 10.1007/978-981-92-1300-9_34.

Yatian Pang, Wenxiao Wang, Francis E. H. Tay, Wei Liu, Yonghong Tian, and Li Yuan. Masked autoencoders for point cloud self-supervised learning. In *Computer Vision – ECCV 2022*, pp. 604–621, 2022.

William Peebles and Saining Xie. Scalable diffusion models with transformers. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 4172–4182, 2023. doi: 10.1109/ICCV51070.2023.00387.

Karsten Roth, Latha Pemula, Joaquin Zepeda, Bernhard Schölkopf, Thomas Brox, and Peter Gehler. Towards total recall in industrial anomaly detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 14318–14328, 2022. doi: 10.1109/CVPR52688.2022.01392.

Lukas Ruff, Jacob R. Kauffmann, Robert A. Vandermeulen, Grégoire Montavon, Wojciech Samek, Marius Kloft, Thomas G. Dietterich, and Klaus-Robert Müller. A unifying review of deep and shallow anomaly detection. *Proceedings of the IEEE*, 109(5):756–795, 2021. doi: 10.1109/JPROC.2021.3052449.

Shunsuke Sakai, Xiangteng He, Chunzhi Gu, Leonid Sigal, and Tatsuhito Hasegawa. InvAD: Inversion-based reconstruction-free anomaly detection with diffusion models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2026.

Jascha Sohl-Dickstein, Eric Weiss, Niru Maheswaranathan, and Surya Ganguli. Deep unsupervised learning using nonequilibrium thermodynamics. In *Proceedings of the International Conference on Machine Learning (ICML)*, volume 37 of *Proceedings of Machine Learning Research*, pp. 2256–2265, 2015.

Jiaming Song, Chenlin Meng, and Stefano Ermon. Denoising diffusion implicit models. In *International Conference on Learning Representations (ICLR)*, 2021.

Yang Song and Stefano Ermon. Generative modeling by estimating gradients of the data distribution. In *Advances in Neural Information Processing Systems*, volume 32, 2019.

Alexandros Stergiou, Brent De Weerdt, and Nikos Deligiannis. Holistic representation learning for multitask trajectory anomaly detection. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, 2024.

Ya Su, Youjian Zhao, Chenhao Niu, Rong Liu, Wei Sun, and Dan Pei. Robust anomaly detection for multivariate time series through stochastic recurrent neural network. In *Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)*, pp. 2828–2837, 2019. doi: 10.1145/3292500.3330672.

Mingxing Tan and Quoc V. Le. EfficientNet: Rethinking model scaling for convolutional neural networks. In *Proceedings of the International Conference on Machine Learning (ICML)*, volume 97 of *Proceedings of Machine Learning Research*, pp. 6105–6114, 2019.

Laurens van der Maaten and Geoffrey Hinton. Visualizing data using t-sne. *Journal of Machine Learning Research*, 9(86):2579–2605, 2008.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. In *Advances in Neural Information Processing Systems*, volume 30, pp. 6000–6010, 2017.

Chengjie Wang, Wenbing Zhu, Bin-Bin Gao, Zhenye Gan, Jiangning Zhang, Zhihao Gu, Shuguang Qian, Mingang Chen, and Lizhuang Ma. 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 (CVPR)*, pp. 22883–22892, 2024.

Chengsen Wang, Zirui Zhuang, Qi Qi, Jingyu Wang, Xingyu Wang, Haifeng Sun, and Jianxin Liao. Drift doesn't matter: Dynamic decomposition with diffusion reconstruction for unstable multivariate time series anomaly detection. In *Advances in Neural Information Processing Systems*, volume 36, 2023.

Hongsong Wang, Andi Xu, Pinle Ding, and Jie Gui. Dual conditioned motion diffusion for pose-based video anomaly detection. *Proceedings of the AAAI Conference on Artificial Intelligence*, 39(7): 7700–7708, 2025. doi: 10.1609/aaai.v39i7.32829.

Renjie Wu and Eamonn J. Keogh. Current time series anomaly detection benchmarks are flawed and are creating the illusion of progress (extended abstract). In *Proceedings of the IEEE International Conference on Data Engineering (ICDE)*, pp. 1479–1480, 2022. doi: 10.1109/ICDE53745.2022.00116.

Julian Wyatt, Adam Leach, Sebastian M. Schmon, and Chris G. Willcocks. AnoDDPM: Anomaly detection with denoising diffusion probabilistic models using simplex noise. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 649–655, 2022. doi: 10.1109/CVPRW56347.2022.00080.

Guoyang Xie, Jinbao Wang, Jiaqi Liu, Jiayi Lyu, Yong Liu, Chengjie Wang, Feng Zheng, and Yaochu Jin. IM-IAD: Industrial image anomaly detection benchmark in manufacturing. *IEEE Transactions on Cybernetics*, 2024.

Enquan Yang, Peng Xing, Hanyang Sun, Wenbo Guo, Yuanwei Ma, Zechao Li, and Dan Zeng. 3CAD: A large-scale real-world 3C product dataset for unsupervised anomaly detection. *Proceedings of the AAAI Conference on Artificial Intelligence*, 39(9):9175–9183, 2025. doi: 10.1609/aaai.v39i9.32993.

Hang Yao, Ming Liu, Zhicun Yin, Zifei Yan, Xiaopeng Hong, and Wangmeng Zuo. GLAD: Towards better reconstruction with global and local adaptive diffusion models for unsupervised anomaly detection. In *Computer Vision – ECCV 2024*, pp. 1–17, 2025. doi: 10.1007/978-3-031-73209-6_1.

Zeqi Ye, Qijie Zhu, Molei Tao, and Minshuo Chen. Provable separations between memorization and generalization in diffusion models. In *International Conference on Learning Representations (ICLR)*, 2026.

Zhiyuan You, Lei Cui, Yujun Shen, Kai Yang, Xin Lu, Yu Zheng, and Xinyi Le. A unified model for multi-class anomaly detection. In *Advances in Neural Information Processing Systems*, volume 35, 2022.

Hui Zhang, Zheng Wang, Dan Zeng, Zuxuan Wu, and Yu-Gang Jiang. DiffusionAD: Norm-guided one-step denoising diffusion for anomaly detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 47(8):7140–7152, 2025. doi: 10.1109/TPAMI.2025.3570494.

Huijie Zhang, Jinfan Zhou, Yifu Lu, Minzhe Guo, Peng Wang, Liyue Shen, and Qing Qu. The emergence of reproducibility and consistency in diffusion models. In *Proceedings of the International Conference on Machine Learning (ICML)*, volume 235 of *Proceedings of Machine Learning Research*, pp. 60558–60590, 2024a.

Jian Zhang, Runwei Ding, Miaoju Ban, and Linhui Dai. PKU-GoodsAD: A supermarket goods dataset for unsupervised anomaly detection and segmentation. *IEEE Robotics and Automation Letters*, 9(3):2008–2015, 2024b. doi: 10.1109/LRA.2024.3352358.

Jiangning Zhang, Haoyang He, Zhenye Gan, Qingdong He, Yuxuan Cai, Zhucun Xue, Yabiao Wang, Chengjie Wang, Lei Xie, and Yong Liu. Ader: A comprehensive benchmark for multi-class visual anomaly detection. *arXiv preprint arXiv:2406.03262*, 2024c.

Baozhu Zhao, Xiaohan Zhang, Jingfeng Guo, and Qi Liu. PointCore: An efficient framework for unsupervised point cloud anomaly detection using joint local-global features. *Neural Networks*, 197:108446, 2026. doi: 10.1016/j.neunet.2025.108446.

Mingyuan Zhou, Huangjie Zheng, Zhendong Wang, Mingzhang Yin, and Hai Huang. Score identity distillation: Exponentially fast distillation of pretrained diffusion models for one-step generation. In *Proceedings of the International Conference on Machine Learning (ICML)*, volume 235 of *Proceedings of Machine Learning Research*, pp. 62307–62331, 2024.

Zheyuan Zhou, Le Wang, Naiyu Fang, Zili Wang, Lemiao Qiu, and Shuyou Zhang. R3D-AD: Reconstruction via diffusion for 3D anomaly detection. In *Computer Vision – ECCV 2024*, pp. 91–107, 2025. doi: 10.1007/978-3-031-72764-1_6.

Jun-Yan Zhu, Philipp Krähenbühl, Eli Shechtman, and Alexei A. Efros. Generative visual manipulation on the natural image manifold. In *Computer Vision – ECCV 2016*, 2016.

Liyun Zhu, Lei Wang, Arjun Raj, Tom Gedeon, and Chen Chen. Advancing video anomaly detection: A concise review and a new dataset. In *Advances in Neural Information Processing Systems*, volume 37, pp. 89943–89977, 2024. doi: 10.52202/079017-2856.

Yang Zou, Jongheon Jeong, Latha Pemula, Dongqing Zhang, and Onkar Dabeer. SPot-the-Difference self-supervised pre-training for anomaly detection and segmentation. In *Computer Vision – ECCV 2022*, 2022. doi: 10.1007/978-3-031-20056-4_23.