A Survey on Diffusion Models for Anomaly Detection

Jing Liu[1,2] , **Zhenchao Ma**[2] , **Zepu Wang**[3] , **Yang Liu**[1,3] , **Zehua Wang**[2] ,
Peng Sun[3] , **Liang Song**[1] , **Bo Hu**[1] , **Azzedine Boukerche**[4] , **Victor C.M. Leung**[2]

[1]Fudan University, Shanghai, China

[2]University of British Columbia, Vancouver, Canada

[3]Duke Kunshan University, Suzhou, China

[4]University of Ottawa, Ottawa, Canada

{jingliu19, yang_liu20, songl, bohu}@fudan.edu.cn, {zhenchaoma, zwang}@ece.ubc.ca,
{zw163, peng.sun568}@duke.edu, aboukerc@uottawa.ca, vleung@ieee.org

Abstract

Diffusion models (DMs) have emerged as a powerful class of generative AI models, showing remarkable potential in anomaly detection (AD) tasks across various domains, such as cybersecurity, fraud detection, healthcare, and manufacturing. The intersection of these two fields, termed diffusion models for anomaly detection (DMAD), offers promising solutions for identifying deviations in increasingly complex and high-dimensional data. In this survey, we review recent advances in DMAD research. We begin by presenting the fundamental concepts of AD and DMs, followed by a comprehensive analysis of classic DM architectures including DDPMs, DDIMs, and Score SDEs. We further categorize existing DMAD methods into reconstruction-based, density-based, and hybrid approaches, providing detailed examinations of their methodological innovations. We also explore the diverse tasks across different data modalities, encompassing image, time series, video, and multimodal data analysis. Furthermore, we discuss critical challenges and emerging research directions, including computational efficiency, model interpretability, robustness enhancement, edge-cloud collaboration, and integration with large language models. The collection of DMAD research papers and resources is available at `https://github.com/fdjingliu/DMAD`.

1 Introduction

The ever-increasing volume, velocity, and variety of data present both opportunities and challenges for anomaly detection (AD). In cybersecurity, real-time threat detection is crucial given the constant influx of network traffic and logs [Jin *et al.*, 2024]. Similarly, robust fraud detection systems are essential for financial institutions processing massive transaction datasets [Wang *et al.*, 2023a]. Healthcare and manufacturing also rely heavily on data for early disease diagnosis and predictive maintenance, respectively [Zhang *et al.*, 2023]. Such applications highlight the growing need for au-

tomated AD methods to efficiently identify outliers in complex datasets. However, traditional techniques, often based on statistical methods or rule-based systems, struggle with the scale and complexity of modern data [Katsuoka *et al.*, 2024]. Specifically, these methods often require extensive manual feature engineering and struggle to adapt to evolving data distributions. Additionally, the sheer data volume can overwhelm traditional methods, rendering them computationally infeasible for real-time applications. Consequently, more sophisticated and scalable AD approaches are needed. For example, weakly supervised approaches leverage limited labeled data to improve performance [Pang *et al.*, 2021], driving innovation toward scalable deep learning architectures that address these fundamental challenges.

Diffusion models (DMs) have emerged as a powerful class of generative models, synthesizing samples across diverse data modalities [Wang *et al.*, 2023c]. Unlike generative adversarial networks (GANs) and variational autoencoders (VAEs), which may exhibit training instability or produce less detailed samples [Dao *et al.*, 2024], DMs generate sharper, more realistic samples through a gradual denoising process [Chen *et al.*, 2024b]. Consequently, DMs excel in capturing a broader range of data distributions compared to GANs [Wyatt *et al.*, 2022], making them particularly suitable for anomaly detection, where their enhanced mode coverage facilitates precise modeling of normal patterns and subsequent identification of distributional outliers [Li *et al.*, 2024].

Recent growth in complex and multi-dimensional data has created new challenges for anomaly detection methods. In this context, DMs have emerged as a promising solution due to their inherent connection to density estimation [Le Lan and Dinh, 2021]. DMs learn the probability distributions of normal data through iterative denoising. Leveraging their exceptional generative abilities, DMs accurately reconstruct normal patterns while capturing underlying manifold structures, enabling anomaly detection through reconstruction error analysis and probability density estimation [Hu and Jin, 2023]. For example, AnoDDPM [Wyatt *et al.*, 2022] uses diffusion models for anomaly detection via reconstruction error, and ODD [Wang *et al.*, 2023b] enhances this by leveraging advanced generative modeling for high-dimensional data. Additionally, the learned score function, representing the gradi-
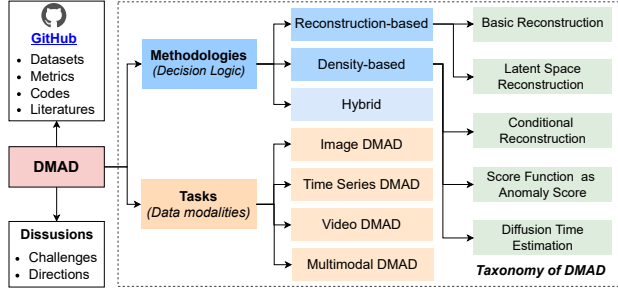
Figure 1: Taxonomy of diffusion models for anomaly detection.

ent of the log-probability density, can be directly used as an anomaly score, as demonstrated in diffusion time estimation [Livernoche *et al.*, 2023].

Motivations. While existing surveys have separately reviewed anomaly detection [Jin *et al.*, 2024; Pang *et al.*, 2021] and diffusion models [Yang *et al.*, 2024b; Luo, 2023], they either focus on traditional deep learning approaches without considering the emerging role of diffusion models, or treat anomaly detection merely as one of many DM applications without in-depth analysis. Recent breakthroughs in diffusion models for anomaly detection (DMAD) demonstrate remarkable potential to revolutionize the field through their unique capabilities in modeling complex data distributions and generating high-quality samples. Despite these advances, there is no existing work has systematically reviewed how DMs enhance anomaly detection or analyzed their emerging role at the intersection of both domains. Our work addresses this gap by providing a timely overview of DMAD, analyzing its theoretical foundations, current achievements, and promising future directions that could fundamentally reshape the landscape of anomaly detection research.

Contributions. The main contributions of this survey are four-fold: (**1**) *A systematic taxonomy*. We present the first comprehensive taxonomy of DMAD (as shown in Fig. 1), categorizing existing methodologies into reconstruction-based, density-based, and hybrid methods, providing a structured framework for understanding this emerging field. (**2**) *A comprehensive review*. Based on the proposed taxonomy, we systematically analyze DMAD across different tasks, including image anomaly detection (IAD), time series anomaly detection (TSAD), video anomaly detection (VAD), and multimodal anomaly detection (MAD), discussing their approaches, strengths, and limitations. (**3**) *Technical challenges and future directions*. We identify and discuss critical challenges in DMAD, including computational costs, interpretability, robustness against adversarial attacks, complex data distributions, edge-cloud collaboration, and integration with large language models (LLMs), while suggesting potential solutions and future research directions. (**4**) *Resources and benchmarks*. We provide a thorough collection of representative methods, implementations, datasets, and evaluation metrics for different tasks, serving as a valuable resource for researchers and practitioners in this field.

2 Preliminaries

This section begins by providing the fundamental concepts of anomaly detection. Following this, we introduce the various types of anomalies, including point, contextual, and collective anomalies. Next, we discuss the basic principles of DMs and their relevance to anomaly detection. We then explore score-based models and score matching techniques. Finally, we present prominent variants of DMs used in anomaly detection, highlighting their advantages and disadvantages.

2.1 Anomaly Detection

Anomaly detection aims to identify instances deviating significantly from the expected data distribution [Wang and Vastola, 2023]. Given a dataset $\mathcal{D} = \{x_1, x_2, ..., x_n\}$, where $x_i \in \mathbb{R}^d$, the goal is to identify the anomalous subset $\mathcal{A} \subset \mathcal{D}$. Anomalous behavior is characterized by low probability density under the typically unknown distribution $p(x)$. Several anomaly types exist, including point anomalies [Hu and Jin, 2023] that are individual deviant instances; contextual anomalies [Sui *et al.*, 2024] that are anomalous within a specific context; and collective anomalies that are groups of instances anomalous only when considered together. Common evaluation metrics include AUC [Xu *et al.*, 2023], measuring the classifier's ability to distinguish between normal and anomalous instances. Additionally, precision, recall, and F1-score assess the trade-off between correct identification and minimizing false alarms [Wang *et al.*, 2023a; Xiao *et al.*, 2023; Zuo *et al.*, 2024]. The choice of metric depends on the application and the relative importance of minimizing false positives and negatives [Pang *et al.*, 2021].

2.2 Diffusion Models

DMs constitute a class of latent variable models defined by forward and reverse Markov processes to corrupt and denoise data, with training based on the variational lower bound of the negative log-likelihood.

Forward Process. Forward process is known as the diffusion process, gradually adds noise to data x_0 over T timesteps, transforming its distribution into an isotropic Gaussian. Specifically, given a variance schedule $\{\beta_t\}_{t=1}^T$, where $0 < \beta_1 < \cdots < \beta_T < 1$, the process generates sequence of latent variables $\{x_1, \ldots, x_T\}$ via a Markov chain with the Gaussian transition kernel:

$$q(x_t|x_{t-1}) = \mathcal{N}(x_t; \sqrt{1 - \beta_t}x_{t-1}, \beta_t\mathbf{I}). \quad (1)$$

Here, $\mathcal{N}(x; \mu, \Sigma)$ denotes a Gaussian distribution with mean μ and covariance Σ, and $\mathbf{I}$ is the identity matrix. The parameter β_t controls the noise added at each step. As t increases, x_T approaches Gaussian noise. Due to the Markov property, the joint distribution is:

$$q(x_{1:T}|x_0) = \prod_{t=1}^{T} q(x_t|x_{t-1}). \quad (2)$$

Consequently, any x_t can be directly sampled from x_0:

$$q(x_t|x_0) = \mathcal{N}(x_t; \sqrt{\bar{\alpha}_t}x_0, (1 - \bar{\alpha}_t)\mathbf{I}), \quad (3)$$

where $\alpha_t = 1 - \beta_t$ and $\bar{\alpha}_t = \prod_{i=1}^{t} \alpha_i$, which provides computational benefits. Related work explores this process through various perspectives, including multiplicative transitions, uniformization techniques for convergence analysis, Gaussian process covariance transformations, tensor network modeling, first-passage time statistics, and infinite-dimensional diffusion processes [Kumar *et al.*, 2023].

Reverse Process. Reverse process aims to learn the conditional probability distribution $p_\theta(x_{t-1}|x_t)$ to recover the original data distribution by iteratively removing noise. It is typically modeled as a Gaussian distribution:

$$p_\theta(x_{t-1}|x_t) = \mathcal{N}(x_{t-1}; \mu_\theta(x_t, t), \Sigma_\theta(x_t, t)), \quad (4)$$

where μ_θ and Σ_θ are the predicted mean and covariance, respectively, parameterized by a neural network (e.g., U-Net [Yan *et al.*, 2024] or Transformer [Cao *et al.*, 2024]) and conditioned on the noisy input x_t at timestep t. While the covariance $\Sigma_\theta(x_t, t)$ can be learned, a common simplification is to use a fixed, time-dependent variance [Salimans and Ho, 2021]. The neural network learns to predict the mean $\mu_\theta(x_t, t)$, which guides denoising. Several parameterization strategies exist, including predicting the mean of $q(x_{t-1}|x_t, x_0)$, added noise ϵ_t, or original data x_0. Starting from Gaussian noise x_T, the reverse process iteratively applies $p_\theta(x_{t-1}|x_t)$ until x_0 is reached.

Training Objective. DMs are trained by maximizing the likelihood of observed data. However, due to the latent variable formulation, direct optimization is intractable. Consequently, training relies on maximizing a variational lower bound (VLB) of the data log-likelihood, derived using variational inference [Bercea *et al.*, 2023] as:

$$\begin{aligned}\mathcal{L}_{\text{vlb}} =\ &\mathbb{E}_q[D_{\text{KL}}(q(x_T|x_0)||p_\theta(x_T))] - \log p_\theta(x_0|x_1) \\ &+ \mathbb{E}_q[\sum_{t=2}^{T} D_{\text{KL}}(q(x_{t-1}|x_t, x_0)||p_\theta(x_{t-1}|x_t))],\end{aligned} \quad (5)$$

where q and p_θ denote the forward and learned reverse process distributions, respectively. Here, D_{KL} is the Kullback-Leibler divergence, x_0 represents the original data, and x_t is the noisy sample at timestep t. A simplified objective minimizes the L_2 distance between true and predicted noise [Han *et al.*, 2024b], can be expressed as:

$$\mathcal{L}_{\text{simple}} = \sum_{t=1}^{T} \mathbb{E}_q\big[||\epsilon_t - \epsilon_\theta(x_t, t)||^2\big], \quad (6)$$

where ϵ_t and $\epsilon_\theta(x_t, t)$ denote the true and predicted noise, respectively. Such simplified objective exhibits computational efficiency and effectiveness [Livernoche *et al.*, 2023]. Additionally, weighting loss terms at different timesteps can improve performance [Choi *et al.*, 2022]. From another perspective, DMs can be viewed as distribution estimators with theoretical guarantees, and the training objective can be based on score matching, where the model estimates the score function [Zhang and Pilanci, 2024]. A unified framework for discrete and continuous-time DMs offers further insights into objective optimization [Lee *et al.*, 2023a].

2.3 Score-based Models and Score Matching

Score Matching for Learning Score Functions. Score-based generative models learn the score function $\nabla_x \log p(x)$, representing the gradient of the log-probability density. By encoding distributional characteristics directly through these gradients, score functions enable powerful generative capabilities without explicit density estimation. Score matching offers a practical way to learn this score function without needing $p(x)$ explicitly. Instead, it minimizes the Fisher divergence between the learned and true score functions [Han *et al.*, 2024b]. Specifically, the score matching objective is $L_{SM}(\theta) = \frac{1}{2}\mathbb{E}_{p(x)}\big[||\nabla_x \log p_\theta(x) - \nabla_x \log p(x)||^2\big]$.

Denoising Score Matching. However, computing the true score is often intractable. Consequently, denoising score matching perturbs data with Gaussian noise and learning the score of this perturbed distribution. The objective becomes $L_{DSM}(\theta) = \frac{1}{2}\mathbb{E}_{q_\sigma(x)}\big[||\nabla_x \log p_\theta(x) - \nabla_x \log q_\sigma(x)||^2\big]$, where $q_\sigma(x)$ is the perturbed distribution. Learning scores at various noise levels allows DMs to capture the data distribution at different scales [Wang *et al.*, 2023d].

Advanced Techniques. Recent research highlights the inherent linear structure within score-based models, particularly at higher noise levels, potentially accelerating sampling [Wang and Vastola, 2023]. Prioritizing specific noise levels during training, based on perceptual relevance, can further improve performance [Choi *et al.*, 2022]. In addition, gradient guidance, incorporating an external objective function's gradient during sampling, adapts pre-trained models to specific tasks [Tur *et al.*, 2023b].

2.4 Variants of Diffusion Models

In this section, we discuss prominent diffusion model variants for anomaly detection, including denoising diffusion probabilistic models (DDPMs), denoising diffusion implicit models (DDIMs), and score-based generative models with stochastic differential equations (Score SDEs), along with their advantages and disadvantages.

DDPMs. DDPMs [Bercea *et al.*, 2023] progressively introduce Gaussian noise to the data across multiple time steps using a Markov chain. A trained neural network then reverses this process, enabling the model to effectively learn complex distributions, which makes it particularly well-suited for anomaly detection [Wyatt *et al.*, 2022]. However, generating samples can be computationally expensive.

DDIMs. DDIMs [Xu *et al.*, 2023] extend DDPMs with a non-Markovian sampling approach, enabling faster generation by skipping steps. Specifically, they modify the reverse process, and by setting a specific parameter to zero, DDIMs allow deterministic sampling, beneficial for consistent reconstructions in anomaly detection [Tebbe and Tayyub, 2024].

Score SDEs. Score SDEs [Deveney *et al.*, 2023] model diffusion as a continuous-time SDE, where the score function, representing the gradient of the log-probability density, can be used for anomaly scoring [Li *et al.*, 2024]. While offering flexibility for complex distributions [Livernoche *et al.*, 2023], training and parameterization choices can be challenging [Graham *et al.*, 2023].
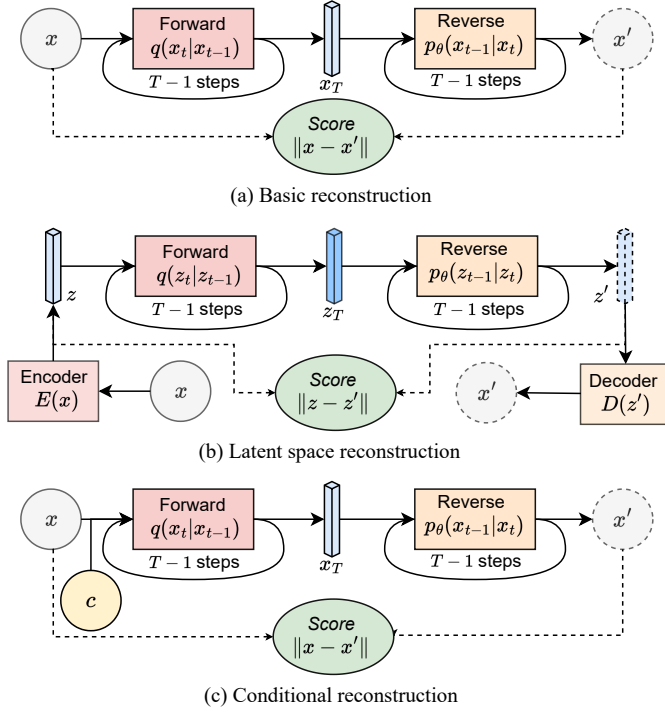
(a) Basic reconstruction

(b) Latent space reconstruction

(c) Conditional reconstruction

Figure 2: Pipeline illustration of reconstruction-based AD methods: (a) basic, (b) latent space, and (c) conditional reconstruction.

3 Methodologies

3.1 Reconstruction-based Anomaly Detection

Reconstruction-based AD (RAD) leverages DMs to identify deviations from learned data distributions. Within this paradigm, we consider three reconstruction strategies, as illustrated in Fig. 2: 1) basic reconstruction, where reconstruction error serves as the anomaly score; 2) latent space reconstruction, incorporating dimensionality reduction for efficiency with high-dimensional data; 3) conditional reconstruction, leveraging auxiliary information like class labels for potentially more refined anomaly detection.

Basic Reconstruction. RAD hinges on the principle that DMs trained on normal data reconstruct normal instances effectively, but struggle with anomalous ones [Hu and Jin, 2023]. Specifically, after training on normal samples, the model reconstructs a test sample by reversing the diffusion process. The difference between the input and its reconstruction constitutes the anomaly score [Zhang *et al.*, 2023]. A higher score indicates a greater likelihood of an anomaly. For instance, GLAD [Yao *et al.*, 2025] employs global and local reconstruction perspectives. In contrast, AutoDDPM [Bercea *et al.*, 2023] focuses on improving normal region reconstruction. Alternatively, some approaches use a learned similarity network for comparison, yielding a semantic distance as the anomaly score [Wang *et al.*, 2023b].

Latent Space Reconstruction. RAD leverages latent space representations for enhanced efficiency and effective handling of high-dimensional data, such as images or videos. Specifically, projecting input data into a lower-dimensional

latent space significantly reduces the computational burden associated with reconstruction. Autoencoders (AEs) are commonly employed for this dimensionality reduction, which learns a compressed representation by encoding data into a latent space and then decoding it back to the original space. Integrating AEs with DMs offers a promising approach. For example, DMs trained on AEs-learned latent representations can capture the normal data distribution in the compressed space [Hu and Jin, 2023]. In addition, NGLS-Diff [Han *et al.*, 2024a] is a novel approach that utilizes DMs within a normal gathering latent space to enhance anomaly detection capabilities for time series data. The latent space reconstruction error then serves as an anomaly score, with deviations indicating potential anomalies [Wang *et al.*, 2023c]. Shared latent spaces across multiple AEs can separate treatment and context. Similarly, supervised dimensionality reduction can optimize the latent space for specific classification tasks, and employing Brownian motion within Diffusion VAEs captures dataset topology [Le Lan and Dinh, 2021].

Conditional Reconstruction. Conditional information, such as class labels, masks, or text descriptions, guides the diffusion model's reconstruction toward expected normal outputs. For example, FDAE [Zhu *et al.*, 2024] utilizes foreground objects and motion information as inputs to train a conditional diffusion autoencoder, enhancing the detection of anomalous regions in an unsupervised manner. Dual conditioning [Zhan *et al.*, 2024] improves multi-class anomaly detection by ensuring prediction and reconstruction accuracy within the expected category. A learnable encoder in RecDMs [Xu *et al.*, 2023] extracts semantic representations for conditional denoising, thus guiding recovery and avoiding trivial reconstructions. Similarly, GLAD [Yao *et al.*, 2025] introduces synthetic anomalies during training to encourage the model to learn complex noise distributions for improved anomaly-free reconstruction. In addition, [Tebbe and Tayyub, 2024] explores dynamic step size computation based on initial anomaly predictions for refined reconstruction. Finally, MDPS [Wu *et al.*, 2024] models normal image reconstruction as multiple diffusion posterior sampling using a masked noisy observation model and a diffusion-based prior, consequently improving reconstruction quality.

3.2 Density-based Anomaly Detection

This section considers two primary methods for density-based AD (DAD), as shown in Fig. 3: Score function and diffusion time estimation (DTE), illustrated by their use of the learned probability density to identify outliers.

Score Function as Anomaly Score. Gradient-based log probability density score functions, $\|\nabla_x \log p(x)\|$, accurately and effectively quantify data point likelihoods within learned distributions [Zhang and Pilanci, 2024]. Within low-density regions, higher score magnitudes reliably indicate potential anomalies by leveraging the score network $s_\theta(x_t, t)$ [Livernoche *et al.*, 2023]. Additionally, probability density characterization through score computation enables direct anomaly assessment without requiring explicit reconstruction or comparative analysis. Theoretical foundations established by [Wang and Vastola, 2023] and [Han *et al.*, 2024b]
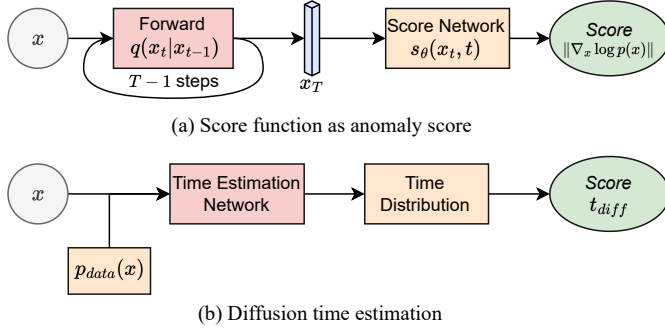
(a) Score function as anomaly score

(b) Diffusion time estimation

Figure 3: Pipeline illustration of density-based AD methods: (a) score function-based and (b) diffusion time estimation method.

Method	DS	Datasets	Results of metrics on datasets
FNDM [2023]	D	BRATS, ISLES	Dice: 76.21% (BRATS), 54.44% (ISLES), VS: 82.28% (BRATS), etc.
ODD [2023b]	D	MNIST, CIFAR-10, etc.	AUROC: 97.8% (MNIST), 84.7% (CIFAR-10), etc.
mDDPM [2024]	D	BraTS21, MSLUB	Dice: 53.02% (BraTS21), 10.71% (MSLUB), etc.
Dif-fuse [2024]	D	IST-3, BraTS, WMH	Dice: 0.699 (BraTS21), 0.569 (WMH), etc.
DIC [2024]	C	VisA, BTAD, MVTec	I-AUROC: 96.0%, P-AUROC: 97.9%, PRO 94.1% on VisA, etc.
DRAD [2024]	C	MVTec-AD	I-AUROC: 99.1% (MVTec-AD), P-AUROC: 97.3% (MVTec-AD), etc.
MDPS [2024]	C	MVTec-AD, BTAD	I-AUROC: 98.4% (MVTec-AD), P-AUROC: 97.0% (MVTec-AD), etc.
DISYRE [2024]	C	CamCAN, ATLAS, BraTS	AP: 0.45 (ATLAS), 0.51 (BraTS-T1), 0.73 (BraTS-T2), etc.

Notes: **Method** highlighted in blue indicates available open-source code links. **DS** denotes diffusion space, where 'D' is discrete and 'C' is continuous. Due to space limitations, only partial datasets and results are shown. For more details, please find our GitHub.

Table 1: Summary of IAD methods with metric results on datasets.

strongly validate this methodology, although recent investigations [Deveney *et al.*, 2023] illuminate potential disparities between SDE and ODE formulations, warranting careful consideration in score-based anomaly detection frameworks.

Diffusion Time Estimation. DTE offers an alternative anomaly detection method leveraging the diffusion process [Livernoche *et al.*, 2023]. Instead of reconstructing the input x, DTE estimates the diffusion timestep t_{diff} required for x to traverse the data distribution $p_{data}(x)$. The estimation is performed by a time estimation network, which predicts the time distribution. Intuitively, a longer t_{diff} suggests an anomaly due to the x's distance from the learned distribution [Luo, 2023]. The estimated diffusion time t_{diff} thus serves as an anomaly score, derived from the mode or mean of its distribution. An analytical form for this density is derivable, and deep neural networks can improve inference efficiency. Consequently, DTE has shown competitive performance, especially in speed, on benchmarks like ADBench, while maintaining or exceeding the accuracy of traditional DDPMs.

3.3 Hybrid Approaches

Hybrid approaches integrate DMs with other anomaly detection techniques to improve performance. For example, RAD can be combined with density estimation, where reconstruction error from DMs is integrated with DAD, like DTE [Livernoche *et al.*, 2023]. Additionally, dynamic step size computation in the forward process, guided by initial anomaly predictions, improves performance by denoising scaled input without added noise [Tebbe and Tayyub, 2024]. In another approach, likelihood maps of potential anomalies from DMs are integrated with the original image via joint noised distribution re-sampling, enhancing healthy tissue restoration [Bercea *et al.*, 2023]. For TSAD, hybrid methods can employ density ratio-based strategies to select normal observations for imputation, combined with denoising diffusion-based imputation to improve missing value generation, especially under anomaly concentration [Xiao *et al.*, 2023]. Advanced reconstruction techniques predict specific denoising steps by analyzing divergences between image and diffusion model priors, supplemented by synthetic abnormal sample generation during training and spatial-adaptive feature fusion during inference [Yao *et al.*, 2025]. Integrating DMs with Transformers for multi-class anomaly detection, where diffusion pro-

vides high-frequency information for refinement, mitigates blurry reconstruction and "identical shortcuts" [Zhan *et al.*, 2024]. Finally, a Bayesian framework employing masked noisy observation models and diffusion-based normal image priors enables effective difference map computation between normal posterior samples and the test image, improving anomaly detection and localization [Wu *et al.*, 2024].

4 Tasks

In this section, we consider four primary tasks of DMAD, as summarized in Tables 1-4, which present representative methods with open-source implementations, datasets, evaluation metrics, and corresponding results. For each task, methods are categorized based on their underlying principles and detection strategies, highlighting the unique challenges and solutions developed for different data modalities.

4.1 Image Anomaly Detection

IAD represents a key application area for DMs, focusing on identifying deviations from normality in both global patterns affecting entire images and local anomalies confined to specific regions. Many approaches leverage DMs' reconstruction capabilities, with methods calculating anomaly scores from reconstruction errors after reconstructing images from noisy inputs [Xu *et al.*, 2023; Tur *et al.*, 2023a; Wang *et al.*, 2023d]. Advanced frameworks like GLAD [Yao *et al.*, 2025] employ global and local adaptive approaches, while ODD [Wang *et al.*, 2023b] introduces similarity networks for semantic distance measurement. Medical imaging applications have shown particular promise, with FNDM [Li *et al.*, 2023] and mDDPM [Iqbal *et al.*, 2024] demonstrating enhanced efficiency through masked diffusion models, while [Fontanella *et al.*, 2024] introduces novel counterfactual generation techniques. Recent research includes AnomalyDiffusion [Hu *et al.*, 2024b] for few-shot generation and [Naval Marimont *et al.*, 2024]'s cold-diffusion approach. MDPS [Wu *et al.*, 2024] and DRAD [Sheng *et al.*, 2024] further enhance detection through masked sampling and noise embedding, while methods like [Tebbe and Tayyub, 2024]

Method	DS	Datasets	Results of metrics on datasets
SimpDM [2024b]	C	Iris, Yacht, etc.	RMSE: 0.059 (Iris), 0.045 (Yacht), 0.102 (Housing), etc.
CoDi [2023a]	C & D	Bank, Heart, etc.	B-F1: 0.4726, B-AUROC: 0.8106, M-F1: 0.62211, etc., on Bank
TimeDiT [2024]	C	Stocks, Energy, ect.	DS: 0.115 (Stocks), 0.1778 (Air Quality), 0.1726 (Energy), etc.
NGLS-Diff [2024a]	C	SWaT, MSL, etc.	F1: 97.52% (SWaT), 90.81% (MSL), 95.30% (SMAP), 98.04% (PSM)
D^3R [2023a]	C	PSM, SMD, SWaT	F1: 76.09% (PSM), 86.82% (SMD), 78.12% (SWaT)
DiffAD [2023]	C	MSL, SWaT, etc.	F1: 94.19% (MSL), 97.66% (SWaT), 97.95% (PSM), 96.95% (SMAP), etc.
DiffADT [2024]	C	SMAP, MSL, SMD	Pre: 96.33% (SMAP), 95.21% (MSL), 79.87% (SMD), etc.

Table 2: Summary of TSAD methods with metric results on datasets.

Method	DS	Datasets	Results of metrics on datasets
MoCoDAD [2023]	D	UB, HR-STC, etc.	AUC: 77.6% (HR-STC), 89.0% (HR-Ave), 68.4% (HR-UB), 68.3% (UB)
[Tur et al., 2023b]	D	UCF, ShT	AUC: 76.36% (ShT), 63.67% (UCF)
DHVAD [2024]	C	Ped2, Ave, ShT	AUC: 98.1% (Ped2), 88.3% (Ave), 78.9% (ShT)
FDAE [2024]	D	Ave, ShT, Ped2	AUC: 97.7% (Ped2), 91.1% (Ave), 73.8% (ShT)
Masked Diffusion [2023]	D	CrossTask,NIV,COIN	SR: 39.17% (CrossTask), 32.35% (NIV), 29.43% (COIN), etc.
FPDM [2023]	D	Ave, ShT, UCF, UB	AUC: 90.1% (Ave), 78.6% (ShT), 74.7% (UCF), 62.7% (UB)
[Wang et al., 2023d]	D	Ped2, Ave, ShT	AUC: 96.5% (Ped2), 92.2% (Ave), 75.4% (ShT)
DiffVAD [2024]	D	Ave, ShT, UCF, etc.	AUC: 81.9% (ShT), 90.3% (Ave), 87.6% (Ped2), etc.
Masked Diffusion [2023]	C	CrossTask,NIV,COIN	SR: 39.17% (CrossTask), 23.47% (CrossTask), 32.35% (NIV), etc.
VADiffusion [2024a]	D	Ped2, Ave, ShT	AUC: 98.2% (Ped2), 94.93% (Ave), 97.32% (ShT)

Table 3: Summary of VAD methods with metric results on datasets.

Method	Mods.	DS	Datasets	Results of metrics on datasets
MPDR [2023]	I, V, A	C	MNIST, CIFAR-10, etc.	AUPR: 0.764 (MNIST), 0.9860 (CIFAR-10), 0.8338 (CIFAR-100), etc.
DIAG [2024]	I, T	C	KSDD2	AP: 80.1% (zero-shot), 92.4% (full-shot)
AnomalyXFusion [2024a]	I, T, Te	C	MVTec-AD, MVTec LOCO	IS: 1.82, IC-LPIPS: 0.33

Notes: **Mods.** indicates modalities, where 'I' is image, 'V' is vector, 'A' is audio, 'T' is text, and 'Te' is texture.

Table 4: Summary of MAD methods with metric results on datasets.

pervised VAD using DMs, relying on high reconstruction error to indicate anomalies, while [Tur et al., 2023b] enhanced this approach by introducing compact motion representations as conditional information. Several innovative architectures have been recently proposed to improve detection accuracy: VADiffusion [Liu et al., 2024a] employs a dual-branch structure combining motion vector reconstruction and I-frame prediction, while FDAE [Zhu et al., 2024] introduces a flow-guided diffusion autoencoder with sample refinement for comprehensive detection of both appearance and motion anomalies. Recent advances include [Yan et al., 2023]'s feature prediction diffusion model and [Cheng et al., 2024]'s denoising diffusion-augmented hybrid framework, both enhancing semantic understanding of normal patterns. Furthermore, [Wang et al., 2023d] proposed an ensemble approach using stochastic reconstructions and motion filters, while [Fang et al., 2023] explored masked diffusion with task-awareness for more focused anomaly detection in specific contexts.

4.4 Multimodal Anomaly Detection

MAD integrates data from various sources to identify deviations from expected behavior. While DMs are relatively nascent in this area, they hold substantial promise through advanced multimodal data integration. For instance, AnomalyXFusion [Hu et al., 2024a] enhances anomaly synthesis by effectively combining image, text, and mask features, while [Flaborea et al., 2023] demonstrates significant success in skeleton-based VAD through motion-conditioned diffusion models. For IAD, combining visual data with text descriptions significantly improves subtle anomaly detection [Capogrosso et al., 2024], with energy-based approaches [Yoon et al., 2023] leveraging manifold structures for more accurate boundary learning. GLAD [Yao et al., 2025] further suggests extending DMs for multimodal data through specialized architectures, while integration with LLMs offers promising directions for context-aware anomaly detection.

5 Discussions

Despite significant progress in DMAD, several critical challenges remain. In this section, we explore key areas that require further research and development.

Computational Cost. Widespread adoption of DMAD remains constrained by substantial computational requirements, particularly when processing high-dimensional data or extended time series. Computational demands manifest during both distribution learning in training and data generation in sampling phases, with the sampling process's iterative nature demanding numerous steps for quality output [Cui et al., 2023]. For example, high-resolution IAD suffers from increased computational overhead due to data volume, model

employ dynamic step size computation to improve localization accuracy.

4.2 Time Series Anomaly Detection

DMs have emerged as a powerful technique for TSAD, effectively learning complex distributions of sequential data [Yang et al., 2024b]. DMs can identify various anomaly types, including point, contextual, and collective anomalies [Sui et al., 2024]. Advanced approaches like NGLS-Diff [Han et al., 2024a] leverage latent spaces, while methods [Zuo et al., 2024] used structured state space layers explicitly to model temporal dependencies. Recently, some innovations address key challenges: D3R [Wang et al., 2023a] tackles non-stationarity through dynamic decomposition, while TimeDiT [Cao et al., 2024] introduces a foundation model approach with multiple masking schemes. Another key challenge is non-stationarity. DiffAD [Xiao et al., 2023] handle missing data through imputation, complemented by self-supervised techniques [Liu et al., 2024b] and co-evolving strategies [Lee et al., 2023a] for mixed-type temporal data.

4.3 Video Anomaly Detection

VAD effectively leverages spatio-temporal information to identify unusual events. DMs have emerged as a highly promising VAD approach due to their ability to learn complex data distributions and generate high-quality reconstructions. For example, [Tur et al., 2023a] investigated unsu-

complexity, and the need to capture temporal dependencies [Wyatt *et al.*, 2022]. However, some researchers are actively exploring solutions. One promising direction is faster sampling methods, such as progressive distillation and optimized sampling schedules like align your steps, which aim to reduce sampling steps while maintaining quality [Salimans and Ho, 2021]. Another approach is model compression through techniques like pruning and quantization. Additionally, efficient architectures, like [Cui *et al.*, 2023] leveraging ER SDEs for faster sampling, and preconditioning methods offer further computational gains.

Interpretability and Explainability. Interpretability remains a key challenge for DMAD [Katsuoka *et al.*, 2024]. Understanding DM's decision-making is crucial, especially in critical applications. Visualizing anomaly scores, like spatially highlighting anomalous image regions as in reconstruction-based methods [Yao *et al.*, 2025], becomes essential. However, the iterative denoising process inherent in DMs complicates interpretability. Explaining anomaly detection requires identifying and explaining deviating features. Integrating DMs with explainable AI (xAI) techniques, such as incorporating attention mechanisms or leveraging LLMs for textual/visual explanations [Wang *et al.*, 2023b], offers a promising research direction.

Complex Data Distributions. A key challenge in applying DMAD lies in handling complex data distributions. Imbalanced datasets, where anomalies are rare, can bias DMs towards the majority class [Yang *et al.*, 2024a], hindering accurate anomaly modeling. Similarly, multimodal datasets, representing distinct normal behaviors, can confound DM learning [Zuo *et al.*, 2024], potentially misclassifying data from less prominent modes. Noisy or missing data further complicates DM training and inference [Choi *et al.*, 2022], as differentiating true anomalies from data imperfections becomes difficult. However, potential solutions exist. Data augmentation techniques can address class imbalance by generating synthetic minority class samples [Yang *et al.*, 2024a]. Robust training methods, such as using semi-unbalanced optimal transport, can enhance DM resilience to noise and outliers [Dao *et al.*, 2024]. Specialized architectures, potentially incorporating mechanisms for handling missing data or modeling multiple modes [Xiao *et al.*, 2023], and dynamic step size computation [Tebbe and Tayyub, 2024] could further improve DM performance on complex distributions.

Robustness and Adversarial Attacks. Adversarial robustness is a critical concern for DMAD. Similar to other deep learning models, DMs are vulnerable to adversarial perturbations, raising concerns about the reliability of diffusion-based AD systems [Chen *et al.*, 2024b]. For example, [Kang *et al.*, 2025] demonstrates the potential of DMs for adversarial defense in classification, directly applying such methods to AD can reduce anomaly detection rates. Specifically, the purification process may remove crucial anomaly signals along with noise. Additionally, adversarial examples exhibit misalignment within DM manifolds, offering a potential detection avenue but requiring further investigation. The observed robustness differences between pixel-space diffusion models (PDMs) and more vulnerable latent diffusion models (LDMs)

[Graham *et al.*, 2023] further underscore the need to consider specific DM types. Consequently, future research should prioritize robust DM-based AD methods, including architectures and training procedures that distinguish genuine anomalies from adversarial noise. Promising directions include defense strategies like DIFFender [Kang *et al.*, 2025], leveraging text-guided diffusion and the adversarial anomaly perception phenomenon, and incorporating insights from adversarial example behavior within DM manifolds.

Edge-Cloud Collaboration. Real-time DMAD faces significant adoption barriers due to their substantial computational demands [Li *et al.*, 2023]. Edge-cloud collaboration offers a promising solution to this challenge by distributing the workload between edge devices and cloud servers. Lightweight DMs deployed on edge devices perform initial anomaly screening, while resource-intensive tasks like full reconstructions are efficiently offloaded to cloud servers [Yan *et al.*, 2024]. In addition, federated learning enables collaborative model training across edge devices and the cloud without sharing sensitive data, thereby enhancing generalization and preserving privacy [Jin *et al.*, 2024]. Dynamic DM partitioning [Chen *et al.*, 2024a] facilitates adaptive resource allocation, optimizing performance based on network conditions and computational demands. Strategic data placement and distributed DNN deployment principles further enhance system efficiency. The integration of proactive detection mechanisms, exemplified by Maat [Lee *et al.*, 2023b], proves particularly valuable for time-critical applications in cloud monitoring and AIOps environments.

Integrating with Large Language Models. Integrating DMs with LLMs offers a highly promising avenue for enhancing anomaly detection, particularly by generating human-interpretable explanations of detected anomalies and incorporating rich contextual information [Kumar *et al.*, 2023]. For example, LLMs can leverage detailed textual descriptions to provide context for observed fluctuations, distinguishing genuine anomalies from expected variations for MAD [Capogrosso *et al.*, 2024]. However, current LLM integration for anomaly detection faces a key challenge: effectively representing and tokenizing temporal data. Existing tokenizers, primarily designed for text, may not adequately capture the subtle nuances of numerical and temporal data, potentially hindering performance [Li *et al.*, 2024]. Consequently, a critical research direction involves developing efficient data representation techniques and specialized LLMs for anomaly detection to fully realize this synergistic approach's potential [Tebbe and Tayyub, 2024].

6 Conclusion

In this survey, we first introduce the core concepts of anomaly detection and diffusion models, providing a foundation for understanding DMAD. Then, we systematically review existing methodologies and their tasks across diverse data types, including image, time series, and multimodal data. Furthermore, we analyze representative approaches for each data type, highlighting their strengths and limitations. To promote further research, we identify key challenges and promising future directions, aiming to advance DMAD research and inspire innovative solutions for real-world applications.

References

[Bercea *et al.*, 2023] Cosmin I. Bercea, Michael Neumayr, Daniel Rueckert, and Julia A. Schnabel. Mask, stitch, and re-sample: Enhancing robustness and generalizability in anomaly detection through automatic diffusion models. *arXiv:2305.19643*, 2023.

[Cao *et al.*, 2024] Defu Cao, Wen Ye, and Yan Liu. Timedit: General-purpose diffusion transformers for time series foundation model. In *ICMLW*, 2024.

[Capogrosso *et al.*, 2024] Luigi Capogrosso, Alvise Vivenza, Andrea Chiarini, Francesco Setti, and Marco Cristani. Exploiting multimodal latent diffusion models for accurate anomaly detection in industry 5.0. In *xAI*, volume 3762, pages 230–235, 2024.

[Chen *et al.*, 2024a] Xiangchun Chen, Jiannong Cao, Yuvraj Sahni, Shan Jiang, and Zhixuan Liang. Dynamic task offloading in edge computing based on dependency-aware reinforcement learning. *IEEE TCC*, 12:594–608, 2024.

[Chen *et al.*, 2024b] Yong Chen, Xuedong Li, Peng Hu, Dezhong Peng, and Xu Wang. Diffilter: Defending against adversarial perturbations with diffusion filter. *IEEE TIFS*, 19:6779–6794, 2024.

[Cheng *et al.*, 2024] Kai Cheng, Yaning Pan, Yang Liu, Xinhua Zeng, and Rui Feng. Denoising diffusion-augmented hybrid video anomaly detection via reconstructing noised frames. In *IJCAI*, volume 2, pages 695–703, 2024.

[Choi *et al.*, 2022] Jooyoung Choi, Jungbeom Lee, Chaehun Shin, Sungwon Kim, Hyunwoo Kim, and Sungroh Yoon. Perception prioritized training of diffusion models. In *CVPR*, pages 11472–11481, 2022.

[Cui *et al.*, 2023] Qinpeng Cui, Xinyi Zhang, Zongqing Lu, and Qingmin Liao. Elucidating the solution space of extended reverse-time sde for diffusion models, 2023.

[Dao *et al.*, 2024] Quan Dao, Binh Ta, Tung Pham, and Anh Tran. A high-quality robust diffusion framework for corrupted dataset, 2024.

[Deveney *et al.*, 2023] Teo Deveney, Jan Stanczuk, Lisa Maria Kreusser, Chris Budd, and Carola-Bibiane Schönlieb. Closing the ode-sde gap in score-based diffusion models through the fokker-planck equation. *arXiv:2311.15996*, 2023.

[Fang *et al.*, 2023] Fen Fang, Yun Liu, Ali Koksal, Qianli Xu, and Joo-Hwee Lim. Masked diffusion with task-awareness for procedure planning in instructional videos. *arXiv:2309.07409*, 2023.

[Flaborea *et al.*, 2023] 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 *ICCV*, pages 1–9, 2023.

[Fontanella *et al.*, 2024] Alessandro Fontanella, Grant Mair, Joanna Wardlaw, Emanuele Trucco, and Amos Storkey. Diffusion models for counterfactual generation and anomaly detection in brain images. *IEEE TMC*, 2024.

[Graham *et al.*, 2023] Mark S. Graham, Walter Hugo Lopez Pinaya, Paul Wright, Petru-Daniel Tudosiu, Yee H. Mah, James T. Teo, H. Rolf Jäger, David Werring, Parashkev Nachev, Sebastien Ourselin, and M. Jorge Cardoso. Unsupervised 3d out-of-distribution detection with latent diffusion models. In *MICCAI*, pages 446–456, 2023.

[Han *et al.*, 2024a] Jiashu Han, Shanshan Feng, Min Zhou, Xinyu Zhang, Yew Soon Ong, and Xutao Li. Diffusion model in normal gathering latent space for time series anomaly detection. In *ECML PKDD*, volume 14943, pages 284–300, 2024.

[Han *et al.*, 2024b] Yinbin Han, Meisam Razaviyayn, and Renyuan Xu. Neural network-based score estimation in diffusion models: Optimization and generalization, 2024.

[Hu and Jin, 2023] Xianyao Hu and Congming Jin. Anodode: Anomaly detection with diffusion ode, 2023.

[Hu *et al.*, 2024a] Jie Hu, Yawen Huang, Yilin Lu, Guoyang Xie, Guannan Jiang, Yefeng Zheng, and Zhichao Lu. Anomalyxfusion: Multi-modal anomaly synthesis with diffusion. *arXiv:2404.19444*, 2024.

[Hu *et al.*, 2024b] Teng Hu, Jiangning Zhang, Ran Yi, Yuzhen Du, Xu Chen, Liang Liu, Yabiao Wang, and Chengjie Wang. Anomalydiffusion: Few-shot anomaly image generation with diffusion model. In *AAAI*, volume 38, pages 8526–8534, 2024.

[Iqbal *et al.*, 2024] Hasan Iqbal, Umar Khalid, Chen Chen, and Jing Hua. Unsupervised anomaly detection in medical images using masked diffusion model. In *MLMI*, pages 372–381, 2024.

[Jin *et al.*, 2024] Ming Jin, Huan Yee Koh, Qingsong Wen, Daniele Zambon, Cesare Alippi, Geoffrey I. Webb, Irwin King, and Shirui Pan. A survey on graph neural networks for time series: Forecasting, classification, imputation, and anomaly detection. *IEEE TPAMI*, 46(12):1–20, 2024.

[Kang *et al.*, 2025] Caixin Kang, Yinpeng Dong, Zhengyi Wang, Shouwei Ruan, Yubo Chen, Hang Su, and Xingxing Wei. Diffender: Diffusion-based adversarial defense against patch attacks. In *ECCV*, pages 130–147, 2025.

[Katsuoka *et al.*, 2024] Teruyuki Katsuoka, Tomohiro Shiraishi, Daiki Miwa, Vo Nguyen Le Duy, and Ichiro Takeuchi. Statistical test on diffusion model-based anomaly detection by selective inference, 2024.

[Kumar *et al.*, 2023] Komal Kumar, Snehashis Chakraborty, and Sudipta Roy. Self-supervised diffusion model for anomaly segmentation in medical imaging. In *PRMI*, volume 14301, pages 359–368, 2023.

[Le Lan and Dinh, 2021] Charline Le Lan and Laurent Dinh. Perfect density models cannot guarantee anomaly detection. *Entropy*, 23:1690, 2021.

[Lee *et al.*, 2023a] Chaejeong Lee, Jayoung Kim, and Noseong Park. Codi: Co-evolving contrastive diffusion models for mixed-type tabular synthesis. In *ICML*, pages 18940–18956, 2023.

[Lee *et al.*, 2023b] Cheryl Lee, Tianyi Yang, Zhuangbin Chen, Yuxin Su, and Michael R. Lyu. Maat: Performance metric anomaly anticipation for cloud services with conditional diffusion. In *ASE*, pages 116–128, 2023.

[Li *et al.*, 2023] Jinpeng Li, Hanqun Cao, Jiaze Wang, Furui Liu, Qi Dou, Guangyong Chen, and Pheng-Ann Heng. Fast non-markovian diffusion model for weakly supervised anomaly detection in brain mr images. In *MICCAI*, volume 14224, pages 579–589, 2023.

[Li *et al.*, 2024] Shu Li, Jiong Yu, Yi Lu, Guangqi Yang, Xusheng Du, and Su Liu. Self-supervised enhanced denoising diffusion for anomaly detection. *Inf. Sci.*, 669:120612, 2024.

[Liu *et al.*, 2024a] Hao Liu, Lijun He, Miao Zhang, and Fan Li. Vadiffusion: Compressed domain information guided conditional diffusion for video anomaly detection. *IEEE TCSVT*, 34:8398–8411, 2024.

[Liu *et al.*, 2024b] Yixin Liu, Thalaiyasingam Ajanthan, Hisham Husain, and Vu Nguyen. Self-supervision improves diffusion models for tabular data imputation. In *CIKM*, pages 1513–1522, 2024.

[Livernoche *et al.*, 2023] Victor Livernoche, Vineet Jain, Yashar Hezaveh, and Siamak Ravanbakhsh. On diffusion modeling for anomaly detection. In *ICLR*, 2023.

[Luo, 2023] Weijian Luo. A comprehensive survey on knowledge distillation of diffusion models, 2023.

[Naval Marimont *et al.*, 2024] Sergio Naval Marimont, Vasilis Siomos, Matthew Baugh, Christos Tzelepis, Bernhard Kainz, and Giacomo Tarroni. Ensembled cold-diffusion restorations for unsupervised anomaly detection. In *MICCAI*, pages 243–253, 2024.

[Pang *et al.*, 2021] Guansong Pang, Chunhua Shen, Longbing Cao, and Anton Van Den Hengel. Deep learning for anomaly detection: A review. *ACM CSUR*, 54(2):38:1–38:38, March 2021.

[Salimans and Ho, 2021] Tim Salimans and Jonathan Ho. Progressive distillation for fast sampling of diffusion models. In *ICLR*, 2021.

[Sheng *et al.*, 2024] Xinyu Sheng, Shande Tuo, and Lu Wang. Surface anomaly detection and localization with diffusion-based reconstruction. In *IJCNN*, pages 1–8, 2024.

[Sui *et al.*, 2024] Jialin Sui, Jinsong Yu, Yue Song, and Jian Zhang. Anomaly detection for telemetry time series using a denoising diffusion probabilistic model. *IEEE Sens. J.*, 24:16429–16439, 2024.

[Tebbe and Tayyub, 2024] Justin Tebbe and Jawad Tayyub. Dynamic addition of noise in a diffusion model for anomaly detection. In *CVPR*, pages 3940–3949, 2024.

[Tur *et al.*, 2023a] Anil Osman Tur, Nicola Dall'Asen, Cigdem Beyan, and Elisa Ricci. Exploring diffusion models for unsupervised video anomaly detection. In *ICIP*, pages 2540–2544, 2023.

[Tur *et al.*, 2023b] Anil Osman Tur, Nicola Dall'Asen, Cigdem Beyan, and Elisa Ricci. Unsupervised video anomaly detection with diffusion models conditioned on compact motion representations. In *ICIAP*, pages 49–62, 2023.

[Wang and Vastola, 2023] Binxu Wang and John J. Vastola. The hidden linear structure in score-based models and its application. *arXiv:2311.10892*, 2023.

[Wang *et al.*, 2023a] 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 *NeurIPS*, volume 36, pages 10758–10774, 2023.

[Wang *et al.*, 2023b] He Wang, Longquan Dai, Jinglin Tong, and Yan Zhai. Odd: One-class anomaly detection via the diffusion model. In *ICIP*, pages 3000–3004, 2023.

[Wang *et al.*, 2023c] Wenjie Wang, Yiyan Xu, Fuli Feng, Xinyu Lin, Xiangnan He, and Tat-Seng Chua. Diffusion recommender model. In *SIGIR*, pages 832–841, 2023.

[Wang *et al.*, 2023d] Zhiqiang Wang, Xiaojing Gu, Jingyu Hu, and Xingsheng Gu. Ensemble anomaly score for video anomaly detection using denoise diffusion model and motion filters. *Neurocomputing*, 553:126589, 2023.

[Wu *et al.*, 2024] Di Wu, Shicai Fan, Xue Zhou, Li Yu, Yuzhong Deng, Jianxiao Zou, and Baihong Lin. Unsupervised anomaly detection via masked diffusion posterior sampling. *arXiv:2404.17900*, 2024.

[Wyatt *et al.*, 2022] Julian Wyatt, Adam Leach, Sebastian M. Schmon, and Chris G. Willcocks. Anoddpm: Anomaly detection with denoising diffusion probabilistic models using simplex noise. In *CVPR*, pages 650–656, 2022.

[Xiao *et al.*, 2023] Chunjing Xiao, Zehua Gou, Wenxin Tai, Kunpeng Zhang, and Fan Zhou. Imputation-based time-series anomaly detection with conditional weight-incremental diffusion models. In *KDD*, pages 1–9, 2023.

[Xu *et al.*, 2023] Haohao Xu, Shuchang Xu, and Wenzhen Yang. Unsupervised industrial anomaly detection with diffusion models. *J. Vis. Commun. Image Represent.*, 97:103983, 2023.

[Yan *et al.*, 2023] Cheng Yan, Shiyu Zhang, Yang Liu, Guansong Pang, and Wenjun Wang. Feature prediction diffusion model for video anomaly detection. In *ICCV*, pages 5504–5514, 2023.

[Yan *et al.*, 2024] Chenqian Yan, Songwei Liu, Hongjian Liu, Xurui Peng, Xiaojian Wang, Fangmin Chen, Lean Fu, and Xing Mei. Hybrid sd: Edge-cloud collaborative inference for stable diffusion models, 2024.

[Yang *et al.*, 2024a] Xiongyan Yang, Tianyi Ye, Xianfeng Yuan, Weijie Zhu, Xiaoxue Mei, and Fengyu Zhou. A novel data augmentation method based on denoising diffusion probabilistic model for fault diagnosis under imbalanced data. *IEEE TII*, 20:7820–7831, 2024.

[Yang *et al.*, 2024b] Yiyuan Yang, Ming Jin, Haomin Wen, Chaoli Zhang, Yuxuan Liang, Lintao Ma, Yi Wang, Chenghao Liu, Bin Yang, Zenglin Xu, Jiang Bian, Shirui Pan, and Qingsong Wen. A survey on diffusion models for time series and spatio-temporal data, 2024.

[Yao *et al.*, 2025] 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 *ECCV*, pages 1–17, 2025.

[Yoon *et al.*, 2023] Sangwoong Yoon, Young-Uk Jin, Yung-Kyun Noh, and Frank Park. Energy-based models for anomaly detection: A manifold diffusion recovery approach. In *NeurIPS*, volume 36, pages 1–21, 2023.

[Zhan *et al.*, 2024] Jiawei Zhan, Jinxiang Lai, Bin-Bin Gao, Jun Liu, Xiaochen Chen, and Chengjie Wang. Enhancing multi-class anomaly detection via diffusion refinement with dual conditioning. *arXiv:2407.01905*, 2024.

[Zhang and Pilanci, 2024] Fangzhao Zhang and Mert Pilanci. Analyzing neural network-based generative diffusion models through convex optimization, 2024.

[Zhang *et al.*, 2023] Hui Zhang, Zheng Wang, Zuxuan Wu, and Yu-Gang Jiang. Diffusionad: Norm-guided one-step denoising diffusion for anomaly detection, 2023.

[Zhang *et al.*, 2024] Menghao Zhang, Jingyu Wang, Qi Qi, Pengfei Ren, Haifeng Sun, Zirui Zhuang, Lei Zhang, and Jianxin Liao. Safeguarding sustainable cities: Unsupervised video anomaly detection through diffusion-based latent pattern learning. In *IJCAI*, volume 8, pages 7572–7580, 2024.

[Zhu *et al.*, 2024] Aoni Zhu, Wenjun Wang, and Cheng Yan. Flow-guided diffusion autoencoder for unsupervised video anomaly detection. In *PRCV*, volume 14430, pages 183–194, 2024.

[Zuo *et al.*, 2024] Haiwei Zuo, Aiqun Zhu, Yanping Zhu, Yinping Liao, Shiman Li, and Yun Chen. Unsupervised diffusion based anomaly detection for time series. *APIN*, 54:8968–8981, 2024.