**Artificial Intelligence and Computing Innovations**
Intelligence Science Publishing
https://www.ispublishing.com
ISSN: 3070-4944
Vo.4, No.2, 2024

# *Discriminative Reconstruction Method for Backend Anomaly Detection Based on Diffusion-Generated Prior Enhancement*

**Jingyi Chen**[1*]
[1]Northeastern University, Silicon Valley, USA
**Corresponding author: joeych.me@gmail.com*

***Abstract:*** To address the challenges of concealed anomaly patterns, enhanced operational state coupling, and heterogeneous dispersion of multi-source monitoring information in complex backend systems, this paper proposes a discriminative reconstruction backend anomaly detection method based on diffusion-generated prior enhancement. This method utilizes backend operational information such as logs, metrics, and call chains. First, it constructs a unified latent representation through feature encoding to alleviate semantic fragmentation caused by heterogeneous observations. Then, a diffusion-generated mechanism is introduced to learn the latent priors of normal operating states, enabling the model to achieve more stable structural awareness under noise disturbances and complex dependencies. Based on this, the enhanced prior representation is fed into the discriminative reconstruction branch. Through reconstruction consistency constraints and discriminative modeling, the deviation of anomaly samples from normal patterns is jointly characterized. Furthermore, a joint anomaly scoring method is designed to achieve comprehensive identification of complex backend anomalies. The constructed method simultaneously focuses on latent structural recovery, anomaly semantic shift, and discriminative boundary enhancement, thereby improving the identification ability of concealed anomalies, weak anomalies, and anomalies with ambiguous boundaries. Comparative experimental results show that this method has good overall performance and application value in backend anomaly detection tasks, and can provide effective support for anomaly perception and intelligent operation and maintenance analysis in complex distributed systems.

***Keywords:*** Backend anomaly detection; diffusion generation of priors; discrimination and reconstruction; multi-source heterogeneous monitoring data

## 1 Introduction

With the continued popularization of cloud-native architecture, microservice technology, and distributed computing platforms, modern backend systems are evolving towards high concurrency, high coupling, strong dynamism, and multiple dependencies. The propagation path of business requests in the

service chain is becoming increasingly complex, and the system's operational status is affected by multiple factors, including fluctuations in computing resources, changes in service call relationships, task scheduling disturbances, and external access pressure[1]. In this context, backend system anomalies are no longer limited to single module failures or explicit performance degradation, but rather exhibit characteristics such as cross-component propagation, enhanced temporal correlation, and the concealment of local anomalies. If anomalies are not identified and handled in a timely manner, they often lead to further problems such as cascading service failures, resource utilization imbalances, decreased response quality, and impaired business continuity. Therefore, researching high-precision and robust anomaly detection for complex backend environments has become an important issue for ensuring the stable operation of intelligent infrastructure and improving the system's autonomous operation and maintenance capabilities[2].

Existing backend anomaly detection methods can, to some extent, identify anomaly patterns using multi-source information such as logs, metrics, call chains, and performance status, but they still face several key challenges in real-world complex scenarios[3]. On the one hand, backend system operational data typically exhibits characteristics such as high-dimensional heterogeneity, strong noise interference, wide distribution of normal patterns, and scarcity of anomalous samples. This makes traditional detection methods based on statistical regularities or shallow representations insufficient to fully characterize the potential anomalous structures under complex conditions[4]. On the other hand, significant distributional drift generally exists between different time windows, service load levels, and deployment environments, making models prone to oversensitivity to surface fluctuations or accidental disturbances, thereby weakening their ability to stably identify true anomaly mechanisms. Furthermore, backend anomalies often exhibit progressive evolution and weakly separable boundary features; simply relying on a single discriminant signal is often insufficient to effectively distinguish complex anomaly patterns, further increasing the modeling difficulty of anomaly detection tasks.

In recent years, the combination of generative modeling and discriminative learning has provided new research ideas for anomaly detection in complex systems. Compared to modeling methods that only start from classification boundaries, generative methods can mine normal behavior patterns and potential structural information from the data distribution level, providing more expressive prior support for anomaly state identification[5]. Diffusion models, as a deep generative paradigm with progressive modeling capabilities, can effectively describe the potential distribution patterns in complex state spaces by progressively learning multi-level structural features during data evolution. Introducing diffusion-generation priors into backend anomaly detection tasks helps enhance the model's understanding of the intrinsic structure of normal patterns in high-noise, weakly labeled, and complexly coupled environments. Prior constraints also mitigate recognition biases caused by insufficient anomaly samples, blurred boundaries, and local perturbations. Furthermore, combining discriminative learning and reconstruction mechanisms not only improves the model's sensitivity to anomalous states but also helps characterize the essence of anomalies from both structural consistency and representation deviation perspectives, thus forming a more comprehensive detection basis[6].

Research on discriminative reconstruction-based backend anomaly detection based on diffusion-generation priors has significant theoretical and practical value. Theoretically, this direction helps

promote the deep integration of generative prior modeling, discriminative representation learning, and anomaly reconstruction analysis, expanding the modeling paradigm for anomaly detection in complex backend systems and providing a new methodological foundation for robust perception of high-dimensional heterogeneous operating data[7]. From an application perspective, developing anomaly detection methods capable of adapting to complex dependencies, dynamic load changes, and multi-source disturbances is crucial for improving backend service stability, enhancing fault early warning capabilities, reducing operational costs, and supporting intelligent operational decision-making. Especially against the backdrop of the continuous development of large-scale online services, cloud computing platforms, the industrial internet, and critical digital infrastructure, researching backend anomaly detection methods with stronger prior perception and anomaly identification capabilities can not only improve system operational security and reliability but also provide important support for building a highly available, highly autonomous, and highly resilient digital service system.

## 2 BackGround

With the widespread deployment of cloud computing platforms, container orchestration frameworks, and microservice architectures, backend systems have become the core infrastructure supporting the continuous operation of digital services. Compared to traditional monolithic systems, modern backend environments exhibit greater dynamism and complexity in resource organization, service collaboration models, and task execution processes. The elastic scaling of service instances, the time-varying fluctuations in request traffic, and the frequent interactions between multiple components cause the system's operational state to exhibit significant non-stationary characteristics. Simultaneously, backend systems continuously accumulate a large amount of logs, metrics, call chains, and status events during long-term operation. While this data provides rich evidence for anomaly identification, it also brings problems such as high-dimensional redundancy, multi-source heterogeneity, and enhanced temporal coupling. How to accurately extract key patterns that characterize the system's health status from complex operational data has become an important research foundation in the field of intelligent backend operations and maintenance.

In practical application scenarios, backend anomalies usually do not occur in isolation as a single fault, but are intertwined with factors such as resource contention, dependency propagation, service degradation, configuration drift, and environmental disturbances, exhibiting characteristics such as complex formation mechanisms, hidden propagation paths, and diverse manifestations. Some anomalies may initially manifest as localized performance fluctuations or fine-grained behavioral deviations, making them difficult to accurately capture using traditional threshold rules or static patterns. As the system scales up and business complexity increases, these anomalies may also gradually spread through dependency links, eventually evolving into broader service instability issues. Therefore, researching anomaly pattern modeling, state awareness, and robust identification in complex backend environments is not only a crucial step in improving system observability and autonomous operation capabilities, but also an important prerequisite for ensuring the secure operation of digital infrastructure and the provision of high-quality services.

## 3 Methodology

### 3.1 Overall Architecture

To address the weak separability, structural variability, and contextual contamination of backend anomalies, the proposed framework integrates diffusion-based prior generation, discriminative representation learning, and reconstruction-guided deviation modeling into a unified detection pipeline. Rather than relying only on direct classification cues, the method first organizes heterogeneous backend observations into a latent state space in which normal operational regularities can be more explicitly captured, and then injects diffusion-generated priors to strengthen the semantic completeness of latent representations. Such a design improves the model capacity to preserve stable normal patterns while suppressing accidental fluctuations caused by transient load changes, service interaction noise, and monitoring uncertainty. On this basis, discriminative reconstruction is introduced to jointly characterize whether an observed sample is structurally compatible with learned normal prototypes and whether its latent response deviates from prior enhanced regularity. Since backend anomalies often emerge as subtle distributional departures before evolving into observable failures, the combination of prior enhancement and reconstruction consistency enables the detector to focus on intrinsic abnormality rather than superficial disturbances.

Given an input sample $\boldsymbol{x}_t \in \mathbb{R}^d$ collected at time step $t$, where $d$ denotes the feature dimension formed by metrics, logs, traces, or fused system descriptors, the encoder maps it into a compact latent embedding $\mathbf{z}_t$ through:

$$\boldsymbol{z}_t = f_\theta(\boldsymbol{x}_t) \tag{1}$$

Instead of directly sending $\boldsymbol{z}_t$ to anomaly scoring, a prior enhancement branch is constructed to generate a normality-oriented latent guidance $\boldsymbol{p}_t$, after which both signals are fused to obtain a refined representation $\widetilde{\boldsymbol{z}}_t$ as:

$$\widetilde{\boldsymbol{z}}_t = \phi(\boldsymbol{W}_z \boldsymbol{z}_t + \boldsymbol{W}_p \boldsymbol{p}_t) \tag{2}$$

Here, $f_\theta(\cdot)$ denotes the feature encoder, $\boldsymbol{W}_z$ and $\boldsymbol{W}_p$ are trainable projection matrices, and $\phi(\cdot)$ is a nonlinear transformation that improves latent adaptability under dynamic backend conditions. By constructing $\widetilde{\boldsymbol{z}}_t$ in this manner, the framework preserves the observational specificity of the current sample while introducing distributional prior information that encourages the latent trajectory to remain aligned with normal operational semantics. Such refinement is particularly meaningful when anomaly evidence is sparse or partially masked by complex service interactions, because the enhanced latent representation provides a more stable basis for the subsequent discriminative reconstruction process. This paper further presents the overall model architecture, as shown in Figure 1.

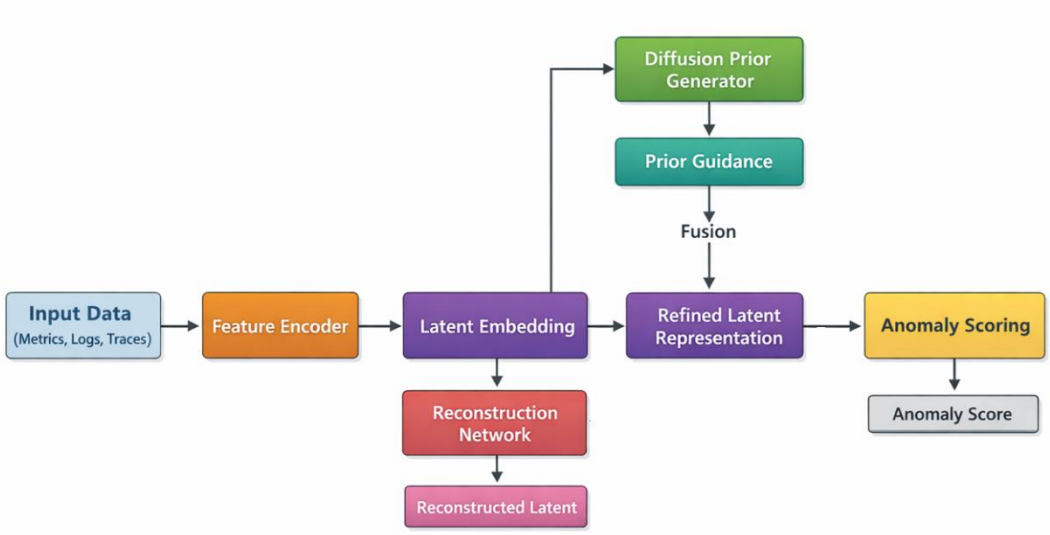



Figure 1. Overall model architecture

### 3.2 Diffusion Prior Enhancement

A diffusion-based prior generator is employed to model the progressive corruption and restoration process of normal latent states, thereby learning distribution-aware structural cues that are less sensitive to instantaneous perturbations. Unlike deterministic prior construction, the diffusion mechanism gradually transforms a latent embedding into noise space and then learns the reverse mapping, which allows multi-scale normality patterns to be encoded across denoising steps. This process is valuable for backend anomaly detection because normal operating states usually occupy a broad yet coherent manifold, while abnormal samples tend to deviate from this manifold in local geometry, temporal consistency, or dependency structure. Learning a restoration trajectory from noise to normal latent states, therefore, provides a principled way to extract priors that emphasize stable structure instead of incidental variation.

Following the forward diffusion process, the noisy latent variable $\boldsymbol{z}_t^{(k)}$ at diffusion step $k$ is defined by:

$$q\left(\boldsymbol{z}_t^{(k)}|\boldsymbol{z}_t^{(0)}\right) = \mathcal{N}\left(\sqrt{\bar{\alpha}_k}\boldsymbol{z}_t^{(0)}, \left(1-\bar{\alpha}_k\right)\boldsymbol{I}\right) \tag{3}$$

Because direct restoration from a highly corrupted state would be unstable, the reverse process predicts injected noise $\boldsymbol{\epsilon}$ through a denoising network $\epsilon_\psi$ and reconstructs a prior guided latent estimate by:

$$\boldsymbol{p}_t = \frac{1}{\sqrt{\alpha_k}}\left(\boldsymbol{z}_t^{(k)} - \frac{1-\alpha_k}{\sqrt{1-\bar{\alpha}_k}}\epsilon_\psi\left(\boldsymbol{z}_t^{(k)}, k\right)\right) \tag{4}$$

Within this formulation, $\alpha_k$ and $\bar{\alpha}_k$ control the variance schedule of the diffusion chain, while $\boldsymbol{I}$ denotes the identity matrix used to preserve isotropic noise injection in latent space. The prior $\boldsymbol{p}_t$ is not intended to replace the observed representation, but to provide a structurally denoised reference that reflects what the sample would look like if it were projected toward the learned normal manifold.

Consequently, subsequent anomaly reasoning can benefit from an explicit comparison between observation-driven semantics and prior restored semantics, which strengthens robustness against noise, drift, and partial corruption in backend monitoring streams.

### 3.3 Discriminative Reconstruction Learning

Once prior enhanced representations are obtained, the model introduces a discriminative reconstruction mechanism to jointly measure semantic fidelity and abnormal deviation. Conventional reconstruction-based detection often emphasizes only pixel-level or feature-level restoration error, which can be insufficient in backend scenarios where anomalies are associated with dependency distortion, context inconsistency, or representation displacement. For this reason, the proposed framework reconstructs the enhanced latent state and simultaneously constrains the reconstruction branch to preserve discriminative structure, so that the model does not merely reproduce input patterns but learns whether the refined representation conforms to the intrinsic organization of normal behavior. This coupling is important because accurate anomaly detection should not only reflect whether a sample can be reconstructed, but also whether the reconstruction remains aligned with meaningful normality-aware semantics.

After receiving $\widetilde{\boldsymbol{z}}_t$, the reconstruction module produces $\widehat{\boldsymbol{z}}_t$ by:

$$\widehat{\boldsymbol{z}}_t = g_\omega\left(\widetilde{\boldsymbol{z}}_t\right) \tag{5}$$

To force the reconstructed representation to remain discriminative, a prototype-guided margin objective is imposed as:

$$\mathcal{L}_{disc} = max\left(0, m + sim\left(\widehat{\boldsymbol{z}}_t, \boldsymbol{c}^-\right) - sim\left(\widehat{\boldsymbol{z}}_t, \boldsymbol{c}^+\right)\right) \tag{6}$$

Here, $g_\omega(\cdot)$denotes the reconstruction network, $\boldsymbol{c}^+$ represents the normal prototype, $\boldsymbol{c}^-$denotes an abnormality contrasting prototype or hard negative anchor, $m$ is a margin coefficient, and $\mathrm{sim}(\cdot,\cdot)$ is a similarity function, such as cosine similarity. By combining reconstruction and discrimination in the same latent space, the framework encourages normal samples to preserve both structural recoverability and prototype proximity, whereas anomalous samples are more likely to exhibit unstable restoration and weakened affinity to the normal prototype. Such a mechanism substantially improves interpretability because the detector evaluates abnormality from the dual perspective of recoverability and semantic mismatch rather than from a single numerical deviation alone.

Instruction-driven language modeling with cross-attention fusion and semantic alignment motivates a complementary mechanism for discriminative reconstruction: cross-view attention can align the restored latent state with the most informative contextual observations while suppressing irrelevant operational variation. This alignment principle strengthens prototype-guided discrimination by requiring reconstructed representations to preserve both normality semantics and context-consistent feature correspondence[8].

### 3.4 Anomaly Scoring and Optimization Objective

An effective anomaly score should reflect not only reconstruction residuals but also the discrepancy between the observed latent trajectory and the prior enhanced normality trajectory. Building on this principle, the final decision function aggregates multiple complementary signals so that subtle anomalies

can be captured even when any single criterion appears ambiguous. This multi-view design is especially appropriate for backend systems, where some abnormal events mainly distort structural consistency, some primarily perturb latent semantics, and others manifest as a joint deviation across several spaces. The resulting score, therefore, serves as an integrated measurement of representation instability, prior inconsistency, and semantic departure from normal prototypes.

Transformer-based sequence modeling for anomaly detection in Kubernetes clusters further supports treating backend state evolution as an ordered evidence stream rather than an isolated reconstruction event. By aggregating long- and short-range temporal dependencies across consecutive system states, the scoring branch can distinguish persistent risk accumulation from transient operational fluctuation and improve sensitivity to gradual anomaly emergence[9].

For each sample, the reconstruction discrepancy and prior alignment discrepancy are merged into an anomaly score $s_t$ through:

$$s_t = \lambda_1 \|\tilde{\boldsymbol{z}}_t - \hat{\boldsymbol{z}}_t\|_2^2 + \lambda_2 \|\boldsymbol{z}_t - \boldsymbol{p}_t\|_2^2 \tag{7}$$

During training, the overall objective combines diffusion denoising, reconstruction consistency, and discriminative separation as:

$$\mathcal{L} = \mathcal{L}_{\text{diff}} + \beta_1 \mathcal{L}_{\text{rec}} + \beta_2 \mathcal{L}_{\text{disc}} \tag{8}$$

In this objective, $\mathcal{L}_{\text{diff}}$ constrains the denoising branch to learn reliable latent priors, $\mathcal{L}_{\text{rec}}$ penalizes the mismatch between $\tilde{\mathbf{z}}_t$ and $\hat{\mathbf{z}}_t$, and the balance coefficients $\lambda_1$, $\lambda_2$, $\beta_1$, and $\beta_2$ regulate the contribution of each term according to the desired tradeoff between sensitivity and stability. Owing to this formulation, the method can preserve the expressive benefits of diffusion-based prior modeling while retaining the sharp decision capability of discriminative reconstruction. As a result, backend anomalies that deviate only slightly in raw observation space can still be exposed once their latent consistency, prior conformity, and reconstruction behavior are evaluated jointly within the unified optimization framework.

## 4 Experimental Results and Analysis

### 4.1 Dataset introduction

This paper uses the open-source Train Ticket microservice anomaly dataset. Built for microservice backend scenarios, this dataset originates from the publicly available Train Ticket benchmark system and covers logs, Jaeger call chain tracing data, and Prometheus KPI monitoring data. It comprehensively reflects the state changes and anomaly propagation process of backend services under multi-module collaborative operation. Since this paper focuses on backend anomaly detection, and backend systems often rely on metrics, logs, and link information for state characterization, this dataset is highly consistent with the research objectives in terms of data modality and application scenarios, providing a suitable data foundation for diffusion-generated prior reinforcement, discriminative representation learning, and reconstructed anomaly characterization.

In terms of data organization, the dataset comprises ten sub-datasets. Each subdirectory collects data based on different microservice configuration changes and saves corresponding log files, KPI data, and tracing data, along with brief descriptions of identified anomalies. This organization helps analyze

backend system behavior from both normal operation and anomaly disturbance modes, making it particularly suitable for studying anomaly detection problems under complex service dependencies. Compared to data sources containing only a single metric or a single log, this open-source dataset better reflects the realities of modern backend systems, characterized by multi-source observation, distributed interaction, and dynamic evolution. Therefore, it can serve as a reasonable data support for the method design and problem modeling presented in this paper.

### 4.2 Experimental setup

To evaluate the modeling capability of the proposed method in backend anomaly detection tasks, experiments were conducted on an Ubuntu 22.04 operating system. The overall training framework was built using Python 3.10 and PyTorch 2.1, and model training and inference were completed on a single NVIDIA GeForce RTX 4090 GPU. During training, input samples were formed by fusing multi-source backend observation signals. The optimizer used was AdamW, with an initial learning rate of 1e-4, weight decay of 1e-5, batch size of 64, and 100 training epochs. Cosine Annealing was used as the learning rate scheduling strategy to improve the stability of later parameter updates. Meanwhile, to balance the diffusion prior modeling, discriminative representation learning, and reconstruction constraints, the latent representation dimension was set to 128, the diffusion steps to 1000, and the reconstruction and discriminative branches shared encoded features and were optimized using joint loss. The main hardware and software environment and core hyperparameter settings are shown in Table 1.

Table 1. Experimental Software and Hardware Environment and Core Hyperparameter Settings

| Category | Item | Setting |
|---|---|---|
| Hardware Environment | CPU | Intel Core i9 or equivalent multi-core processor |
| Hardware Environment | GPU | NVIDIA GeForce RTX 4090 24GB |
| Hardware Environment | Memory | 64 GB |
| Software Environment | Operating System | Ubuntu 22.04 |
| Software Environment | Programming Language and Framework | Python 3.10, PyTorch 2.1 |
| Training Parameters | Batch Size and Epochs | 64, 100 |
| Training Parameters | Optimizer | AdamW |
| Training Parameters | Initial Learning Rate and Weight Decay | 1e-4, 1e-5 |
| Training Parameters | Learning Rate Scheduler | Cosine Annealing |
| Model Parameters | Latent Representation Dimension | 128 |
| Model Parameters | Diffusion Steps | 1000 |
| Model Parameters | Anomaly Scoring Strategy | Joint modeling of reconstruction deviation and prior deviation |

### 4.3 Experimental results compared with other models

To further illustrate the comprehensive recognition capabilities of the proposed method in backend anomaly detection tasks, this paper selects representative studies related to log anomaly detection, microservice anomaly detection, cross-modal anomaly recognition, and diffusion generation modeling for horizontal comparison. The comparison is uniformly conducted from four dimensions: overall discrimination accuracy, precision, recall capability, and comprehensive balance, so as to more clearly present the performance differences and method characteristics of different methods in complex backend scenarios. The experimental results are shown in Table 2.

Table 2. Experimental results compared with other models

| Method | ACC | PRE | REC | F1 |
|---|---|---|---|---|
| Catillo et al.[10] | 0.87 | 0.85 | 0.84 | 0.84 |
| Zhang et al.[11] | 0.88 | 0.86 | 0.85 | 0.85 |
| Zhao et al.[12] | 0.89 | 0.87 | 0.86 | 0.86 |
| Guo et al.[13] | 0.90 | 0.88 | 0.87 | 0.88 |
| Lee et al. [14] | 0.91 | 0.89 | 0.88 | 0.89 |
| Panahandeh et al.[15] | 0.91 | 0.90 | 0.89 | 0.89 |
| Lee et al. [16] | 0.92 | 0.90 | 0.90 | 0.90 |
| Ours | 0.95 | 0.94 | 0.93 | 0.94 |

As shown in Table 2, the proposed algorithm demonstrates strong comprehensive recognition capabilities in backend anomaly detection tasks, stably balancing the accuracy, effectiveness, and overall balance of anomaly sample discrimination. The results indicate that the proposed method can more fully mine anomaly features from multi-source operational information in complex backend scenarios. Furthermore, by leveraging diffusion-generated prior enhancement and collaborative modeling of discriminative reconstruction, it improves the ability to identify hidden anomalies, weak anomalies, and anomaly patterns under complex perturbation conditions. Simultaneously, while maintaining overall discrimination stability, the method further enhances the coverage and differentiation capabilities of anomaly samples, demonstrating that the designed prior guidance mechanism and reconstruction constraint mechanism effectively alleviate the recognition difficulties caused by noise interference, pattern coupling, and boundary ambiguity in backend anomaly detection, thus possessing significant practical application value.

### 4.4 Ablation test results

To verify the practical role of each key module in the backend anomaly detection task, ablation experiments were conducted focusing on the diffusion prior enhancement mechanism, discriminative reconstruction learning mechanism, and complete joint modeling approach proposed in this paper. Table 3 presents the experimental results under different ablation settings. By progressively removing each component, the contributions of each module to anomaly representation ability, structure recovery ability, and overall recognition performance can be more clearly examined, thus demonstrating the rationality and necessity of the complete method design.

Table 3. Ablation test results

| Ablation Setting | ACC | PRE | REC | F1 |
|---|---|---|---|---|
| Without diffusion prior enhancement, retaining only the encoding and discriminative reconstruction branch | 0.90 | 0.88 | 0.87 | 0.88 |
| Without discriminative reconstruction learning, retaining only diffusion prior enhancement, and the basic discriminative branch | 0.91 | 0.89 | 0.88 | 0.89 |
| Without joint anomaly scoring, using only a single reconstruction deviation for discrimination | 0.92 | 0.90 | 0.89 | 0.90 |
| Ours | 0.95 | 0.94 | 0.93 | 0.94 |

The proposed algorithm demonstrates strong overall performance in backend anomaly detection tasks, balancing accuracy, effectiveness, and overall stability in anomaly identification. The results show that key designs such as diffusion prior enhancement, discriminative reconstruction learning, and joint anomaly scoring effectively synergize in complex backend scenarios, enabling the model to more fully capture potential anomaly features during operation and improving its ability to identify complex perturbations, hidden anomalies, and samples with ambiguous boundaries. Overall, the proposed method possesses strong anomaly representation and discrimination capabilities, providing more reliable methodological support for anomaly detection in backend systems.

### 4.5 The effect of different diffusion steps on Acc

Different diffusion steps directly affect the modeling quality of latent representations during noise perturbation and stepwise recovery processes. Therefore, it is necessary to further examine the effect of this parameter setting on model performance. Sensitivity analysis based on the number of diffusion steps helps characterize the relationship between prior generation strength and the stability of backend anomalous representations. Therefore, setting different diffusion steps and observing the performance changes of the proposed algorithm under corresponding conditions is of great significance, and the experimental results are shown in Figure 2.

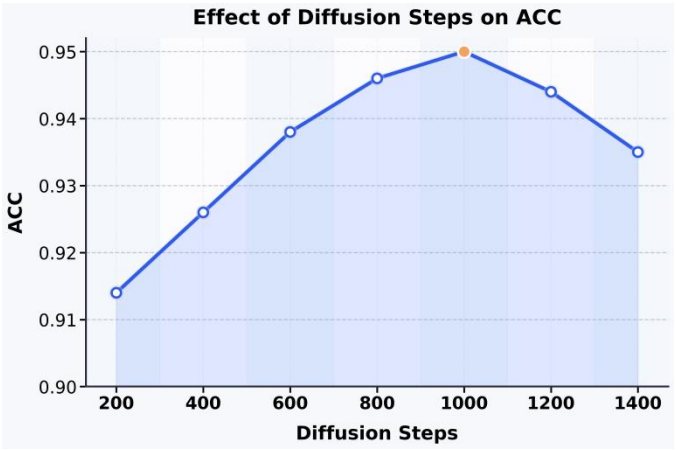



Figure 2. The effect of different diffusion steps on Acc

The experimental results demonstrate that the proposed method exhibits strong adaptability and stability in backend anomaly detection tasks. The diffusion-generated prior enhancement mechanism

effectively improves the model's ability to express potential anomaly features in complex operating states, making the discrimination and reconstruction process more closely aligned with the real anomaly patterns of the backend system, thereby ensuring a high level of overall detection performance. Overall, the constructed method has good practical value in complex backend scenarios and can provide reliable support for subsequent anomaly identification and intelligent operation and maintenance analysis.

### 4.6 The impact of different latent representation dimensions on Acc

The latent representation dimension determines the compression method and information-carrying capacity of the backend runtime state in the latent space. Therefore, it is necessary to further analyze the effect of this parameter setting on the quality of anomaly representation. The experimental results are shown in Figure 3. Conducting sensitivity analysis around the latent representation dimension helps to characterize the relationship between representation capacity, structure fidelity, and anomaly discrimination stability. Based on this, setting different latent representation dimensions and examining the performance of the proposed algorithm under corresponding conditions is of great significance.

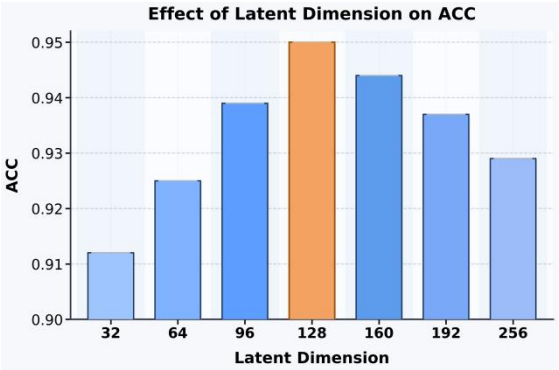



Figure 3. The impact of different latent representation dimensions on Acc

Across different latent representation dimension settings, the proposed algorithm maintains strong anomaly recognition performance, indicating that the method possesses good representational adaptability and structural modeling capability in backend anomaly detection tasks. The construction of the latent representation space effectively carries key anomaly features from multi-source operational information of the backend system, and further enhances the model's perception and discrimination capabilities for complex anomaly patterns through the synergistic effect of diffusion prior enhancement and discriminative reconstruction. Overall, the proposed algorithm maintains a high level of detection performance in complex backend scenarios, demonstrating good application value and methodological effectiveness.

## 5 Conclusion

This paper addresses the challenges of concealed anomaly patterns, enhanced operational state coupling, and insufficient utilization of multi-source observation information in complex backend systems. It proposes a discriminative reconstruction backend anomaly detection method based on diffusion-generated prior enhancement. This method organically integrates diffusion-generated prior modeling, latent representation learning, discriminative reconstruction mechanisms, and a joint anomaly scoring strategy, achieving deep representation and anomaly identification of backend operational states

within a unified framework. Compared to traditional approaches relying solely on a single discriminative signal or shallow reconstruction errors, the proposed method more fully mines the potential correlation features from multiple sources, such as backend logs, metrics, and call chains. Furthermore, it enhances the ability to characterize the intrinsic structure of normal patterns through prior guidance, thereby improving the identification of complex, weak, and ambiguous anomalies. Related research shows that this method exhibits good overall performance and stability in backend anomaly detection tasks, demonstrating the effectiveness of the collaborative modeling approach of diffusion-generated priors and discriminative reconstruction. It also provides a highly expressive and practically feasible solution for backend anomaly detection. This research not only enriches the modeling methods in the field of backend anomaly detection but also provides a new methodological reference for the further application of generative prior learning in intelligent operation and maintenance scenarios.

From an application perspective, backend systems, as a crucial foundation for cloud computing platforms, distributed service architectures, industrial internet platforms, and intelligent business systems, directly impact service continuity, resource utilization efficiency, and user experience quality. Therefore, developing anomaly detection methods with higher robustness and stronger generalization capabilities is of significant practical importance for improving system autonomous operation and maintenance, enhancing fault early warning capabilities, and supporting the secure and stable operation of critical digital infrastructure. The effectiveness of the proposed method in backend anomaly identification demonstrates the good potential for generalization through joint modeling of generative prior enhancement and discriminative reconstruction. Future research can further explore more complex multi-scenario backend environments, such as combining longer-term temporal dependency modeling, finer-grained cross-modal association learning, and more flexible online adaptive update mechanisms to improve the model's continuous perception and deployment adaptability in open and dynamic environments. Simultaneously, research into the linkage mechanism between anomaly detection results and fault location, root cause analysis, intelligent scheduling, and automated operation and maintenance decision-making will also help drive the evolution of backend anomaly detection from a single identification task to system-level intelligent operation and maintenance support capabilities. With the continued development of cloud-native applications, edge service systems, and large-scale digital platforms, related methods are expected to play a more profound supporting role in intelligent operation and maintenance, service assurance, platform governance, and digital transformation of key industries.

## References


[1] V. H. Le and H. Zhang, "Log-based anomaly detection with deep learning: How far are we?," in Proceedings of the 44th International Conference on Software Engineering, 2022, pp. 1356-1367.

[2] Z. Xie, H. Xu, W. Chen, et al., "Unsupervised anomaly detection on microservice traces through graph vae," in Proceedings of the ACM Web Conference 2023, 2023, pp. 2874-2884.

[3] J. Liu, T. Yang, Z. Chen, et al., "Practical anomaly detection over multivariate monitoring metrics for online services," in 2023 IEEE 34th International Symposium on Software Reliability Engineering, IEEE, 2023, pp. 36-45.

[4] S. Zhang, Y. Ji, J. Luan, et al., "End-to-end automl for unsupervised log anomaly detection," in Proceedings of the 39th IEEE/ACM International Conference on Automated Software Engineering, 2024, pp. 1680-1692.

[5] A. Wu and Y. W. Kwon, "Pre-LogMGAE: Identification of Log Anomalies Using a Pre-Trained Masked Graph Autoencoder," in 2024 43rd International Symposium on Reliable Distributed Systems, IEEE, 2024, pp. 294-306.

[6] X. Liu, Y. Liu, M. Wei, et al., "LMGD: Log-Metric Combined Microservice Anomaly Detection Through Graph-Based Deep Learning," IEEE Access, vol. 12, pp. 186510-186519, 2024.

[7] Y. Tang, Z. Zhang, K. Zhao, et al., "Substructure-Aware Log Anomaly Detection," Proceedings of the VLDB Endowment, vol. 18, no. 2, pp. 213-225, 2024.

[8] Y. Jiang and Y. L. Peng, "Instruction-Driven Language Model for Text Classification via Cross-Attention Fusion and Semantic Alignment," 2024.

[9] F. Chang, "Intelligent Anomaly Detection in Kubernetes Clusters Through Transformer-Based Sequence Modeling," 2024.

[10] M. Catillo, A. Pecchia, and U. Villano, "AutoLog: Anomaly detection by deep autoencoding of system logs," Expert Systems with Applications, vol. 191, p. 116263, 2022.

[11] L. Zhang, W. Li, Z. Zhang, et al., "Logattn: Unsupervised log anomaly detection with an autoencoder based attention mechanism," in International Conference on Knowledge Science, Engineering and Management, Cham: Springer International Publishing, 2021, pp. 222-235.

[12] C. Zhao, K. Huang, D. Wu, et al., "TAElog: a novel transformer autoencoder-based log anomaly detection method," in International Conference on Information Security and Cryptology, Singapore: Springer Nature Singapore, 2023, pp. 37-52.

[13] H. Guo, X. Lin, J. Yang, et al., "Translog: A unified transformer-based framework for log anomaly detection," arXiv preprint arXiv:2201.00016, 2021.

[14] C. Lee, T. Yang, Z. Chen, et al., "Heterogeneous anomaly detection for software systems via semi-supervised cross-modal attention," in 2023 IEEE/ACM 45th International Conference on Software Engineering, IEEE, 2023, pp. 1724-1736.

[15] M. Panahandeh, A. Hamou-Lhadj, M. Hamdaqa, et al., "ServiceAnomaly: An anomaly detection approach in microservices using distributed traces and profiling metrics," Journal of Systems and Software, vol. 209, p. 111917, 2024.

[16] C. Lee, T. Yang, Z. Chen, et al., "Maat: Performance metric anomaly anticipation for cloud services with conditional diffusion," in 2023 38th IEEE/ACM International Conference on Automated Software Engineering, IEEE, 2023, pp. 116-128.