

# Deep Learning Advancements in Anomaly Detection: A Comprehensive Survey

Haoqi Huang, Ping Wang<sup>✉</sup>, *Fellow, IEEE*, Jianhua Pei<sup>✉</sup>, *Graduate Student Member, IEEE*, Jiacheng Wang<sup>✉</sup>, Shahen Alexanian, and Dusit Niyato<sup>✉</sup>, *Fellow, IEEE*

**Abstract**—The rapid expansion of data from diverse sources has made anomaly detection (AD) increasingly essential for identifying unexpected observations that may signal system failures, security breaches, or fraud. As datasets become more complex and high-dimensional, traditional detection methods struggle to effectively capture intricate patterns. Advances in deep learning have made AD methods more powerful and adaptable, improving their ability to handle high-dimensional and unstructured data. This survey provides a comprehensive review of over 180 recent studies, focusing on deep learning-based AD techniques. We categorize and analyze these methods into reconstruction-based and prediction-based approaches, highlighting their effectiveness in modeling complex data distributions. Additionally, we explore the integration of traditional and deep learning methods, highlighting how hybrid approaches combine the interpretability of traditional techniques with the flexibility of deep learning to enhance detection accuracy and model transparency. Finally, we identify open issues and propose future research directions to advance the field of AD. This review bridges gaps in existing literature and serves as a valuable resource for researchers and practitioners seeking to enhance AD techniques using deep learning.

**Index Terms**—Anomaly detection, deep learning, data reconstruction and prediction, Internet of things, comprehensive survey.

## I. INTRODUCTION

A n anomaly refers to an observation that significantly deviates from the expected behavior in a system, often appearing unusual, inconsistent, or unexpected [1]. Despite the fact that outliers typically constitute only a small fraction of a dataset, they are often highly crucial because they carry important information and can reveal critical insights during analysis. Consequently, anomaly detection (AD) is the process of identifying such anomalous observations using various methods and algorithms, which aids decision-makers in better understanding data patterns and behaviors.

The rapid development of the Internet of Things (IoT) has revolutionized the way data is generated, collected, and analyzed across various domains. IoT systems leverage a wide array of interconnected sensors and devices to collect massive

H. Huang, P. Wang and S. Alexanian are with the Lassonde School of Engineering, York University, Toronto, ON M3J 1P3, Canada (e-mail: joycehq@yorku.ca; pingw@yorku.ca; yu263319@my.yorku.ca).

J. Pei is with the State Key Laboratory of Advanced Electromagnetic Technology, School of Electrical and Electronic Engineering, Huazhong University of Science and Technology, Wuhan 430074, China (e-mail: jianhuapei@hust.edu.cn).

J. Wang and D. Niyato are with the School of Computer Science and Engineering, Nanyang Technological University, Singapore (e-mail: jcwang\_cq@foxmail.com; dniyato@ntu.edu.sg).

amounts of real-time data in diverse applications, including smart cities [2], industrial automation [3], healthcare [4], and transportation [5], etc. This proliferation of sensor data introduces unprecedented opportunities for enhancing operational efficiency and decision-making processes. However, it also presents significant challenges, as the data is often high-dimensional, noisy, and prone to anomalies caused by faulty sensors, environmental changes, or malicious attacks [6]. Detecting anomalies in data is critical for ensuring system reliability, security, and performance [7].

AD methodologies can be systematically classified according to various criteria. One prominent classification framework differentiates these methods into supervised, semi-supervised, and unsupervised approaches, predicated on the availability and nature of labeled data [8]. Supervised learning-based AD algorithms necessitate a fully labeled dataset, where each data point is explicitly annotated as either normal or anomalous. This labeling process facilitates the model's ability to discern and learn the underlying characteristics that differentiate anomalous instances from normal ones, thereby enhancing its detection accuracy. Semi-supervised learning-based methods, on the other hand, operate with a dataset comprising a substantial volume of unlabeled data alongside a smaller subset of labeled instances. These labels may include both normal and anomalous data, or in certain cases, solely normal instances [9]. In scenarios where only normal data is labeled, the semi-supervised approach converges towards unsupervised methodologies, as the model predominantly learns normal behavior patterns and identifies anomalies as deviations from these learned patterns. Unsupervised learning-based AD methods eschew the need for labeled data entirely, leveraging the intrinsic structural properties of the dataset to autonomously identify anomalies [10] [11]. In practical applications, a significant portion of contemporary AD research gravitates towards unsupervised methods [12]. This preference is largely driven by the substantial imbalance between the number of normal instances and anomalies, which complicates the acquisition of a sufficiently large labeled dataset required for effective supervised learning [13]. Moreover, anomalies are frequently correlated with critical failures or hazardous events, rendering the labeling process both costly and logically challenging. Another key classification criterion is the nature of the dataset, particularly whether it comprises time-series data, which distinguishes AD methods into time-series [14] and non-temporal approaches. The applications of time-series and non-temporal AD will be discussed in detail in Section III.

In addition to the temporal aspect, AD techniques can also

be categorized based on their underlying paradigms: traditional methods and deep learning-based methods. Traditional techniques encompass statistical approaches [15], distance-based methods [16], and clustering algorithms [17]. These approaches generally rely on estimating the probability distribution of normal data to predict anomalies. However, since the early 20th century, the fields of data science, machine learning, deep learning, and artificial intelligence have witnessed exponential growth, with significant implications for AD [18]. Particularly in recent years, the advent of soft-computing techniques has significantly influenced the development of deep learning-based methods. These techniques are characterized by their ability to handle imprecise, uncertain, and nonlinear data, making them highly suitable for applications involving deep learning. Consequently, deep learning-based methods have been propelled to the forefront due to their superior capability to learn expressive representations of complex data, including high-dimensional, temporal, spatial, and graph-structured data [19]. By proficiently modeling intricate patterns and relationships inherent in the data, deep learning approaches have proven remarkably effective in identifying anomalies across a wide range of challenging and complex datasets. This paper concentrates specifically on AD methods based on deep learning models, with the objective of providing a comprehensive review of this rapidly evolving field.

#### A. Contrasting Traditional Models with Deep Learning Models

Traditional AD methods [20], such as statistical techniques, clustering algorithms [21], and Principal Component Analysis (PCA) [22], have long been established as reliable tools across a wide spectrum of applications due to their simplicity, interpretability, and low computational overhead. These characteristics make them particularly promising in scenarios where model transparency and efficiency are paramount. Statistical techniques, for example, provide clear, rule-based mechanisms for detecting anomalies, while clustering algorithms are effective in grouping similar data points and isolating outliers in relatively low-dimensional datasets. Similarly, PCA has been widely adopted for dimensionality reduction, enabling effective AD by isolating principal components that capture major variations in the data [17]. Despite these advantages, traditional methods often encounter significant limitations when applied to modern, complex datasets. Statistical techniques generally assume that data adheres to specific distributions. However, this assumption is rarely met in real-world scenarios, where data often exhibits non-Gaussian distributions and heavy tails. Clustering-based methods, while useful in many contexts, check to accurately define clusters, particularly when anomalies do not present clear separability from normal data. PCA, on the other hand, relies heavily on the assumption of linearity and extensive feature engineering, making it less effective at capturing the nuanced, non-linear patterns prevalent in high-dimensional datasets [22]. These constraints have prompted a shift towards more advanced approaches capable of handling the increasing complexity of modern data.

In contrast, deep learning models have recently emerged as a powerful alternative, addressing many of the shortcom-

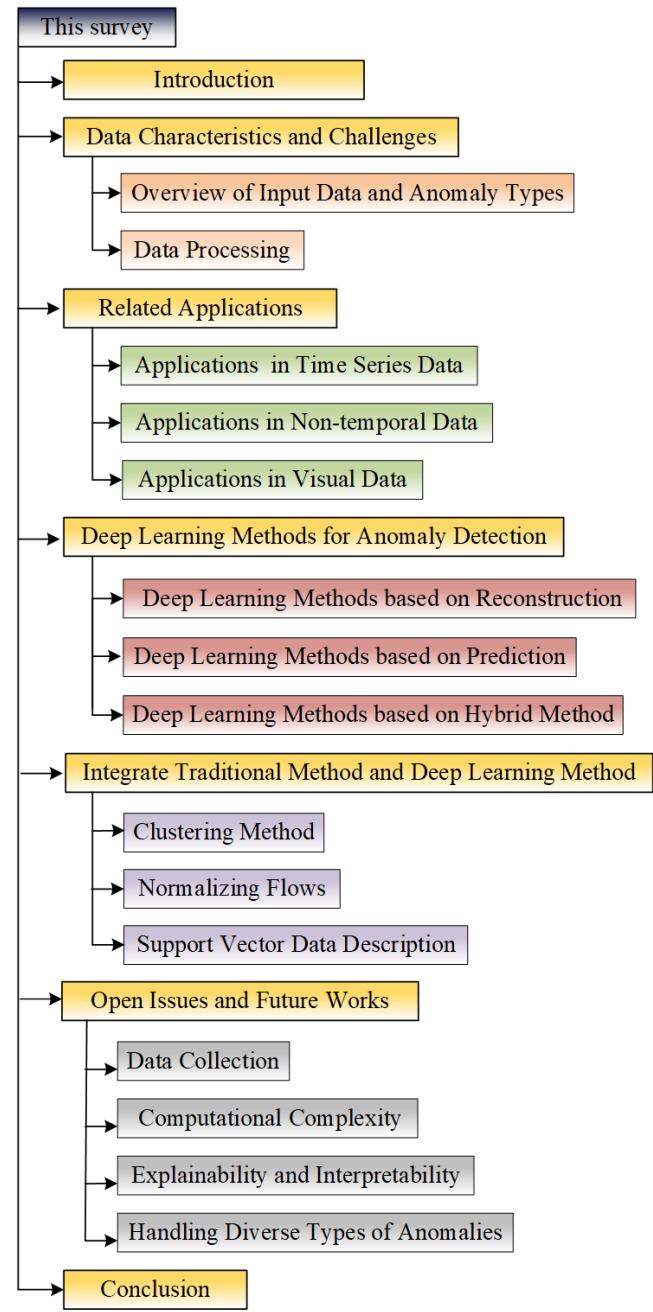


Fig. 1. The anatomy of this survey.

ings inherent in traditional approaches. Deep neural networks (DNNs) possess the capacity to autonomously learn complex patterns and hierarchical representations from raw data, thereby obviating the need for labor-intensive feature engineering [23]. This characteristic is particularly advantageous in the detection of subtle and multifaceted anomalies that might elude traditional methods [24]. By leveraging their multi-layered architectures, deep learning models excel in processing high-dimensional and unstructured data, such as images, videos, and text, which are often challenging for conventional methods to handle effectively [25]. These models are adept at capturing non-linear relationships and interactions within the data, offering a more flexible and robust framework for AD

[26]. Consequently, there has been a significant shift away from purely traditional AD techniques towards the adoption of deep learning methodologies.

Nonetheless, it is crucial to acknowledge that traditional AD models retain certain advantages, notably in their simplicity, interpretability, and lower computational overheads [27]. These characteristics make them particularly appealing in scenarios where model transparency and computational efficiency are crucial. In recognition of these strengths, Section V of this paper will introduce and discuss various existing approaches that integrate traditional methods with deep learning techniques. These hybrid methods aim to leverage the strengths of both paradigms, resulting in more robust and efficient AD systems.

### B. Comparison With Existing Surveys

In recent years, the field of AD has seen a surge in research, particularly with the advent of deep learning methods. Numerous surveys have been published, each attempting to provide a comprehensive overview of the field. However, many of these surveys focus on broader historical developments or cover deep learning techniques only up to a certain point in time. For example, surveys such as [19], [28], [29], and [23] primarily cover techniques developed up to 2020. While these surveys are valuable, they do not reflect the most recent advancements in the field. Furthermore, specific models such as Generative Adversarial Network (GAN)-based AD have been explored in-depth by studies [30], [31], [32], [33], and [34]. However, these studies primarily address foundational approaches and lack coverage of advanced techniques like conditional GANs, cycle-consistent GANs, and GANs integrated with self-supervised learning. Emerging hybrid models, combining GANs with Variational Autoencoders (VAEs) or autoencoders for improved robustness, are also underrepresented. In contrast, our survey covers the literature from 2019 to 2024, providing a timely and comprehensive overview of the latest advancements. By focusing on recent trends and evolving techniques, including enhanced architectures and hybrid frameworks, our work offers a more current perspective, bridging existing gaps and guiding future research directions in AD.

### C. Contributions and Structure

This survey systematically reviews over 160 recent research papers on AD, including publications from leading journals (IEEE, ACM, Springer, Elsevier) and top-tier conferences (AAAI, CCS, ICCV) spanning from 2019 to 2024. By focusing on cutting-edge advancements in deep learning-based methods, this survey ensures a comprehensive and up-to-date overview of the field. The contributions of this survey are summarized as follows:

- This survey addresses gaps in prior surveys by highlighting advanced techniques that were previously underexplored, including conditional GANs, cycle-consistent GANs, and hybrid frameworks combining GANs with VAEs. These models are introduced and analyzed to demonstrate their strengths and weaknesses.

- This survey provides a detailed comparison of reconstruction-based and prediction-based methods. To enhance clarity and usability, we summarize key strengths, weaknesses, and applications in structured tables, offering readers insights into the trade-offs of different models.
- Recognizing the strengths of traditional methods, this survey explores their integration with deep learning models. Hybrid approaches, such as clustering, normalizing flows, and support vector data descriptions combined with deep learning, are analyzed to address complex challenges in AD.

The organization of this survey is shown in Fig.1. Section II provides an overview of data characteristics and anomaly types, followed by a discussion of common data processing challenges and mitigation strategies critical to effective AD. Section III explores the related applications of AD. Section IV categorizes and analyzes deep learning methods for AD, highlighting their effectiveness and limitations. Section V discusses the integration of traditional methods with deep learning, including clustering methods, normalizing flows, and support vector data descriptions. Section VI highlights open issues and future directions, such as challenges in data collection, computational complexity, explainability, and handling diverse anomaly types. Finally, Section VII concludes the survey with a summary and potential directions for future research.

## II. DATA CHARACTERISTICS AND CHALLENGES

### A. Overview of Input Data and Anomaly Types

In AD, input data presents unique challenges due to its structure, dimensionality, and temporal nature. Different types of data require specialized techniques to effectively identify anomalies, and the nature of anomalies themselves can vary greatly depending on the domain and data format [28]. For instance, visual data such as images and videos may exhibit anomalies associated with spatial or temporal inconsistencies, while time series data often involves anomalies related to trends or sudden changes in values over time. To better understand these variations, we first categorize data into textual, audio, image, and video formats, highlighting their respective characteristics and the challenges they pose for AD. Beyond this classification, data can also be viewed through the lens of temporal dependencies, distinguishing between time-series data, which captures sequential patterns over time, and non-temporal data, where observations are independent of temporal order. This dual perspective provides a comprehensive framework for analyzing how different types of anomalies manifest across various data formats. Furthermore, the nature of anomalies themselves can vary depending on the data format. Point anomalies, sequence anomalies, and outliers may all manifest differently across different data types and structures. Understanding these distinctions is essential for selecting the appropriate AD techniques [29], as a deep understanding of data characteristics and anomaly types ensures that detection methods are effectively tailored to capture the specific behaviors and patterns indicative of anomalies.

### 1) Categorization by Data Type:

- **Textual Data:** Textual data consists of sequences of discrete symbols, such as characters, words, or phrases, structured in a linear format. Unlike other data types, textual data conveys information through syntactic and semantic relationships. It can be found in various forms, including documents, chat messages, emails, and system logs. Anomalies in textual data may appear as irregular word sequences, syntactic inconsistencies, missing or misplaced words, or semantically incoherent phrases.
- **Audio Data:** Audio data captures variations in amplitude and frequency over time, representing spoken language, environmental sounds, or machine signals. It can be stored as waveforms or transformed into frequency-domain representations like spectrograms. Unlike textual data, audio data is continuous and often requires spectral analysis to extract meaningful patterns. Anomalies in audio data manifest as unexpected distortions, unusual frequency shifts, missing segments, or abnormal sound patterns caused by malfunctioning equipment, altered speech, or environmental noise.
- **Image Data:** Image data consists of two-dimensional pixel grids, where each pixel represents intensity or color information. Unlike sequential data, image data encodes spatial relationships, capturing textures, shapes, and patterns. Image anomalies often appear as distortions, irregular textures, missing components, or unexpected objects that deviate from normal patterns. For instance, these can result from manufacturing defects, medical imaging errors, or environmental changes in satellite imagery.
- **Video Data:** Video data extends image data by incorporating a temporal dimension, forming sequences of frames over time. Each frame within a video is an image, and the relationships between frames capture motion and dynamic interactions [35]. Unlike static images, video data requires modeling temporal dependencies, making AD more complex. Anomalies in video data include irregular movements, unexpected scene transitions, or unusual object behaviors, which are commonly observed in surveillance footage, traffic monitoring, and activity recognition.
- **Tabular Data:** Tabular data consists of structured records organized in rows and columns, where each row represents an entity or event, and each column corresponds to an attribute. This type of data is widely used in databases, spreadsheets, financial records, and sensor logs. Unlike the other data types, tabular data can contain numerical, categorical, or mixed-format information. Anomalies in tabular data include missing values, unexpected categorical labels, numerical outliers, or inconsistent relationships between attributes.

### 2) Categorization by Temporal Characteristics:

- **Time-based data:** Time-based data can be represented as a sequence of observations recorded over time, and it may consist of either a single variable (univariate) or multiple variables (multivariate). We can generalize the

representation of both univariate and multivariate time series using the following formula:  $X = \{x_{t,j}\}_{t \in T, j \in J}$ , where  $t \in T$  denotes the time index, with  $t$  representing a specific time step and  $T$  being the set of all time steps in the dataset. Similarly,  $j \in J$  represents the dimension or variable index, where  $j$  refers to a particular variable and  $J$  is the set of all variables or dimensions in the data. When  $|J| = 1$ , the series is univariate, meaning there is only one variable observed over time. In contrast, when  $|J| > 1$ , the series is multivariate, indicating that multiple variables are recorded simultaneously at each time step. Each observation  $x_{t,j}$  corresponds to the value of the  $j$ -th variable at time  $t$ . Among the five previously introduced data types, audio, video, and certain types of textual and tabular data are inherently time-based. Audio data is naturally sequential, with sound signals evolving over time, making anomalies such as distortions or frequency shifts dependent on temporal patterns. Video data extends image sequences over time, requiring the detection of abnormal object movements, scene transitions, or motion inconsistencies. Textual data, such as streaming logs, system event records, or chat conversations, also exhibits temporal dependencies, where anomalies may appear as unexpected event sequences or irregular timing between log entries. Similarly, tabular data in the form of financial transactions, sensor readings, or stock prices follows a time-series format, where anomalies may indicate fraud, equipment failure, or unusual market behaviors.

- **Non-temporal data:** Non-temporal data refers to observations that lack a temporal sequence, where the relationships between data points are independent of time. Such data is prevalent across industries that rely on static datasets or event-based observations. AD in non-temporal data focuses on identifying irregularities by analyzing data characteristics, patterns, or statistical properties rather than temporal dependencies. This process is crucial for uncovering hidden risks, fraudulent activities, or system malfunctions in contexts where time is not a defining factor. Among the five data types, image and certain types of tabular data are the most common forms of non-temporal data. Image data, such as medical scans, industrial defect detection images, or satellite photos, captures spatial relationships but does not depend on a temporal sequence. Anomalies in such data typically appear as unusual textures, distortions, or unexpected objects. Tabular data, when not used for time-series analysis, is also non-temporal, such as customer records, product attributes, or static financial datasets. In these cases, AD focuses on identifying outliers, inconsistencies, or unusual relationships between different features rather than changes over time.

### 3) Types of Anomalies:

- **Point Anomalies:** A single data point deviates significantly from the expected behavior in the dataset. These are common across both time-based and non-time-based data, representing sudden outliers or unusual values.
- **Contextual Anomalies:** A data point is considered

anomalous only when it is analyzed within a specific context or surrounding data. In time-based data, this could involve seasonal trends or time-of-day variations, whereas in non-time-based data, it could depend on relationships between variables.

- **Subsequence Anomalies:** A contiguous sequence of data points behaves abnormally, typically found in time series data. These anomalies are significant when the temporal order of data points plays a key role in detecting deviations from expected patterns.
- **Cluster-based and Correlation Anomalies:** Anomalies that occur when a group of data points, or relationships between variables, deviate from expected patterns. This is more prominent in non-time-based data, where detecting irregular clusters or correlations between features is essential for AD.

## B. Data Processing

Effective AD requires careful preparation and preprocessing of input data to ensure that detection algorithms can operate effectively. In many cases, raw data contains inherent challenges that can significantly hinder the performance of AD models. These challenges arise from the complexity of real-world data, including high dimensionality, missing or sparse values, skewed class distributions, and noise that can obscure true anomalies. Without addressing these issues, AD methods may struggle to accurately identify rare or subtle deviations in the data, leading to false positives, missed anomalies, or inefficient computations. Therefore, appropriate data preprocessing steps are crucial for improving detection accuracy, robustness, and overall system reliability. This subsection outlines some of the most common data processing issues and their implications for AD, along with strategies to mitigate these challenges.

1) *Dimensionality:* High-dimensional data makes AD more complex due to the “curse of dimensionality”. As datasets expand in size and complexity—particularly with the rise of “big data”, characterized by large-scale, high-velocity data generated from diverse sources, it becomes increasingly difficult for AD methods to maintain accuracy [36]. To address this issue, dimensionality reduction is a common approach that transforms a large set of input features into a smaller, more focused feature set [37]. While traditional methods such as PCA [38] are frequently used, they may struggle to capture non-linear relationships in complex data. For instance, Sakurada *et al.* [39] compare autoencoders, which perform non-linear dimensionality reduction, with linear PCA and kernel PCA on both synthetic and real-world datasets. The study reveals that on the nonlinear and high-dimensional synthetic Lorenz dataset, AE achieved a relative AUC improvement of 26.83% compared to linear PCA. This highlights that autoencoders can even detect anomalies in data with relatively high intrinsic dimensionality, where linear PCA struggles to perform.

2) *Sparsity:* Sparse data, where many values are missing or incomplete, poses significant challenges for AD. Sparse datasets can lead to reduced detection accuracy, as missing or incomplete data points may obscure the underlying patterns necessary for detecting anomalies [36]. Cheng *et al.*

[40] highlight that in high-dimensional settings, the sparsity problem is further amplified as the data becomes more spread out, increasing the risk of missing critical information that signals anomalies. To address these challenges, Li *et al.* [41] propose an improved low-rank and sparse decomposition model (LSDM) for hyperspectral AD. Their approach models sparse components as a Gaussian Mixture (MoG), effectively capturing anomalous patterns within complex datasets by leveraging the low-rank structure. In contrast, Han *et al.* [42] take a different approach by introducing sparse autoencoders to learn sparse latent representations from high-dimensional input data. Through experiments on three real-world cyber-physical system datasets, the study shows that mining sparse latent patterns from high-dimensional time series can significantly improve the robustness of AD models.

3) *Class Imbalance:* In most AD tasks, the occurrence of anomalies is significantly rarer than normal data points, resulting in a class imbalance problem. This imbalance can cause detection algorithms to be overly biased toward the majority class (normal data), leading to a higher rate of false negatives where critical anomalies are missed. In imbalanced datasets, it is often possible to achieve an overall high accuracy, while the recall score for the minority class (anomalies) remains very low [43]. Traditional methods to mitigate this issue involve oversampling the minority class or undersampling the majority class [44]. Recent research has increasingly focused on introducing Data Generation Models (DGM) to improve the representation of the minority class in AD. For instance, Dlamini *et al.* [45] use Conditional Generative Adversarial Networks (CGANs) to generate synthetic samples for the minority class and combines this with KL divergence to guide the model in accurately learning the distribution of the minority class.

4) *Noise in Data:* Noise refers to random or irrelevant information present in the data, which can obscure true anomalies and lead to false positives. In addition, during the training process of AD models, the high complexity of the model and the presence of noisy data can lead to overfitting, where the model inadvertently learns to fit the reconstruction error from noisy inputs rather than focusing on genuine anomalies [46]. To reduce the impact of noisy data, Zhang *et al.* [47] incorporate a Maximum Mean Discrepancy (MMD) to encourage the distribution of low-dimensional representations to approximate a target distribution. The goal is to align the distribution of noisy data with that of normal training data, thereby reducing the risk of overfitting. Furthermore, Chen *et al.* [48] propose a novel method called Noise Modulated Adversarial Learning, where noise images from a predefined normal distribution are fed into the discriminator network as negative samples. This adversarial process modulates the training of the reconstruction network, balancing the learning between the two networks to improve robustness against noise.

5) *Privacy of data:* In many fields, such as healthcare, finance, and cybersecurity, data used for AD often contains sensitive or personal information. Ensuring the privacy and security of this data is paramount, as improper handling could lead to serious legal and ethical violations. Hassan *et al.* [49] conduct an in-depth investigation into the privacy of AD

models in blockchain technology. To address these privacy concerns, Federated Learning (FL), a distributed machine learning paradigm, has emerged as a promising supplement to AD [50]. FL allows distributed clients to collaboratively train a shared model while protecting the privacy of their local data. For example, Idrissi *et al.* [51] propose Fed-ANIDS, which leverages FL to address the privacy issues associated with centralized Network Intrusion Detection Systems (NIDS). This model was applied to various settings and popular datasets, demonstrating its ability to achieve high performance while preserving the privacy of distributed client data. Cui *et al.* [52] further introduce GAN into FL and design a new algorithm model that injects controllable noise into local model parameters, ensuring both AD utility and compliance with differential privacy requirements.

### III. RELATED APPLICATIONS

With the rapid advancement of deep learning models, AD has become more efficient and adaptable. These sophisticated models have been widely applied across various domains, enhancing the ability to identify irregular patterns in complex and high-dimensional datasets. In the previous chapter, we categorized data based on temporal characteristics into time-series and non-time-series data. However, visual data presents unique challenges, detection requirements, and a wide range of applications, making it difficult to be strictly classified as either time-series or non-time-series data. It can be static (e.g., images) or dynamic (e.g., videos), where images are typically considered non-time-series data, while videos fall under time-series data. Visual data is extensively used in fields such as medical imaging, autonomous systems, and surveillance, where detecting anomalies requires specialized deep learning techniques that differ from traditional numerical or categorical data analysis. To better reflect its broad applications and distinct computational needs, we discuss visual data separately. Based on this classification, we will now explore the applications of deep learning in AD from three perspectives: time-series data, non-temporal data, and visual data.

#### A. Applications in Time Series Data

Time series data, defined by its sequential nature over time, is fundamental to many systems where the temporal order of events critically influences analysis and decision-making processes. AD in time series data has become an indispensable technique across various industries, enabling the early detection of irregular patterns that may indicate underlying issues or emerging threats. The applications of time series AD are extensive, impacting critical areas such as traffic monitoring, power system management, and healthcare. In the following sections, we present how these applications leverage AD to enhance operational efficiency, ensure system reliability, and improve safety across these fields.

*1) Traffic Monitoring:* Time series AD plays a pivotal role in modern traffic management systems. As demonstrated in [53], real-time data from loop detection sensors are integrated and analyzed to predict traffic volume and enhance system safety. The ability to detect anomalies in traffic patterns is

essential for anticipating and responding to potential incidents before they escalate. For instance, Li *et al.* [54] present a method that identifies traffic incidents by detecting anomalies in traffic time series data, thereby helping users avoid accidents and reduce travel time. Furthermore, high-speed driving is identified as a significant contributor to traffic accidents [55]. By monitoring and analyzing sudden increases in vehicle speed, AD techniques can predict and prevent accidents more effectively, providing a critical tool for improving road safety. Zhao *et al.* [56] further validate the efficacy of unsupervised AD methods in assessing elevated road traffic accident risks, specifically by analyzing volume and speed data from traffic on Yan'an elevated road. This approach enhances the ability to detect and respond to hazardous traffic conditions in real-time, underscoring the indispensable role of AD in traffic management.

*2) Power System:* AD is a vital element in ensuring the stability, security, and reliability of electrical grids. By continuously monitoring grid data, these techniques can swiftly identify deviations from normal operational patterns, which may indicate issues such as natural faults or malicious cyber-attacks. The ability to detect these anomalies in real-time is crucial for preventing potential outages and maintaining a consistent power supply. For instance, Li *et al.* [57] highlight that accurate and real-time AD can enhance grid stability by over 20%, providing rapid response capabilities that significantly bolster the system's defense against both natural disruptions and cyber threats. Furthermore, the introduction of a residential electrical load AD framework, as demonstrated in [58], has been shown to significantly improve both load prediction accuracy and AD, thereby optimizing demand-side management (DSM) in residential areas. In terms of cybersecurity, the MENSA Intrusion Detection System (IDS) [59] has proven to be a formidable tool in smart grid environments, effectively detecting operational anomalies and classifying a wide range of cyberattacks. This capability not only protects critical infrastructure but also underscores the indispensable role of AD in modern power system management.

*3) Healthcare:* AD plays a crucial role in healthcare by enabling continuous monitoring of patient vital signs, such as heart rate and blood pressure, to swiftly identify abnormal conditions that may require urgent medical intervention. The application of AD in medical signal analysis is particularly important, as highlighted in [60], where the identification of data samples that deviate from the typical data distribution can reveal underlying issues such as noise, changes in a patient's condition, or the emergence of new and previously undetected medical conditions. This capability is essential for ensuring accurate diagnosis and timely patient care. Furthermore, Keeley *et al.* [61] demonstrate that AD algorithms can effectively identify irregularities in heart rate data, which not only facilitates faster emergency responses but also provides deeper insights into a patient's health status. This, in turn, enhances overall patient care while also reducing the cognitive load on healthcare professionals by automating the detection of potential issues.

### B. Applications in Non-temporal Data

AD in non-temporal data plays a critical role in ensuring operational integrity, security, and financial stability. By focusing on identifying irregularities within independent events or static datasets, it addresses potential risks such as fraud, system failures, and malicious activities. Unlike time-series applications, non-temporal AD leverages data patterns and statistical analysis to uncover deviations that signal anomalies. In the following, we present specific applications across domains such as finance and cybersecurity, showcasing its significant impact on safeguarding critical systems and operations.

1) *Finance*: In the financial sector, non-temporal data AD is pivotal for identifying fraudulent transactions, credit scoring anomalies, and unusual trading activities. Unlike time series data, these financial fraud detection tasks often involve independent events, such as individual transactions or credit score evaluations, which do not rely on temporal sequences. Instead, the focus is on transaction characteristics and patterns that may indicate fraudulent behavior. Various data mining techniques, including SVM, Naïve Bayes, and Random Forest, are extensively employed to detect different forms of financial fraud, such as bank fraud, insurance fraud, financial statement fraud, and cryptocurrency fraud [62]. As highlighted by [63], AD is critical in quickly identifying activities that deviate from normal patterns, thereby enabling rapid intervention to minimize financial losses.

2) *Cybersecurity*: AD is a fundamental component of maintaining a secure and resilient cyberspace. As [64] points out, advanced security controls and resilience analysis are crucial during the early stages of system deployment to ensure long-term sustainability. AD plays a pivotal role in this process by identifying unauthorized access, malicious activities, and network intrusions that deviate from established norms. This capability is essential for safeguarding network security and preventing potential breaches. Early research in deep learning-based network intrusion detection focused on architectures such as Autoencoders (AE), Deep Belief Networks (DBN), and Recurrent Neural Networks (RNN) [24]. As deep learning technology has advanced, more sophisticated models have been developed for detecting anomalies in cybersecurity. For instance, Singh *et al.* [65] illustrate the benefits of AD in wide-area protection schemes (WAPS) by using a deep learning-based cyber-physical AD system (CPADS) to detect and mitigate data integrity and communication failure attacks in centralized Remedial Action Schemes (CRAS). Similarly, Nagarajan *et al.* [66] highlights the effectiveness of AD in enhancing the security of Cyber-Physical Systems (CPSs) by accurately identifying anomalous behaviors, thereby addressing the growing challenges posed by sophisticated cyber-attacks and the increasing volume of data.

### C. Applications in Visual data

AD in visual data, encompassing images and videos, plays a vital role in numerous industries where visual inspection is critical. Applications range from detecting defects in manufacturing processes to identifying medical abnormalities in

imaging, monitoring public safety through surveillance systems, and ensuring quality control in production lines. By leveraging advanced deep learning techniques, AD methods can automatically identify and analyze irregularities with high precision, reducing reliance on manual inspection and improving efficiency. In this section, we explore key applications of visual data-based AD, highlighting its transformative impact across various domains.

1) *Medical Imaging*: AD in medical imaging is indispensable across numerous medical specialties, playing a crucial role in the early detection and diagnosis of diseases. In radiology, it is employed to identify anomalies in X-rays [67], brain imaging [68], and CT scans [69], thereby aiding in the accurate diagnosis of various conditions. However, as [70] highlights, anomalies in medical images often closely resemble normal tissue, posing a significant challenge to detection due to their subtle differences. This similarity requires the use of sophisticated techniques to effectively distinguish between normal and anomalous data. For example, Draelos *et al.* [71] demonstrate the power of machine learning in radiology, significantly enhancing the classification performance for multiple abnormalities in chest CT volumes, achieving an AUROC greater than 0.90 for 18 different abnormalities. Additionally, Shvetsova *et al.* [72] showcase a novel method for AD in medical images, which dramatically improves the detection of subtle abnormalities in complex, high-resolution images, such as chest X-rays and pathology slides—scenarios where traditional models often fail. Furthermore, Zhao *et al.* [73] introduce the SALAD framework, which enhances AD in medical images by utilizing self-supervised and translation-consistent features from normal data. This approach is particularly effective in situations where labeled anomalous images are scarce, thereby improving detection accuracy in challenging medical imaging tasks.

2) *Video Monitoring*: Video AD (VAD) has become increasingly crucial with the rise of large-scale multimedia data analysis, particularly in the processing of video data [74]. VAD focuses on identifying unusual patterns or behaviors in video footage that deviate from the norm, making it a vital tool in several domains. In security and surveillance, VAD is used to monitor public spaces, buildings, and secure areas, enabling the detection of suspicious activities, unauthorized access, and unusual crowd behaviors, thereby enhancing public safety [75]. In the realm of traffic monitoring, VAD facilitates the real-time identification of accidents and irregular traffic patterns, allowing for prompt response and management [76]. Additionally, VAD is applied in behavioral analysis to detect abnormal behaviors in various environments, such as schools, workplaces, and public transportation systems, contributing to the maintenance of safety and order. For example, Chen *et al.* [77] propose a bidirectional prediction framework specifically designed for AD in surveillance videos. This innovative approach employs forward and backward prediction subnetworks to predict the same target frame, constructing a loss function based on the real target frame and its bidirectional predictions. Experimental results demonstrate that this model outperforms existing approaches on various surveillance video datasets, including those featuring pedestrians and street scenes, showcas-

ing its superior performance in accurately detecting anomalies in real-world surveillance scenarios.

#### IV. DEEP LEARNING METHODS FOR ANOMALY DETECTION

The application of deep learning to AD has revolutionized the way we identify irregularities in both time-based and non-time-based datasets [78]. Traditional methods, such as statistical analysis and clustering, have been commonly used to detect anomalies. However, these methods often struggle with high-dimensional data, complex relationships, and capturing intricate patterns. Deep learning models, with their ability to learn hierarchical representations and detect subtle anomalies, have emerged as powerful tools to overcome these limitations.

As shown in Fig.2, this section introduces three major deep learning approaches applied to AD: reconstruction-based methods, prediction-based methods, and hybrid approaches. Each approach leverages the strengths of deep learning in distinct ways to improve AD accuracy, particularly in scenarios where data patterns are complex, unstructured, or temporal.

##### A. Deep learning methods for Anomaly Detection based on Reconstruction

Reconstruction-based approaches operate by training a model to learn the underlying distribution of normal data [79]. Once trained, the model attempts to reconstruct incoming data. The reconstruction error, which is the difference between the original data and its reconstruction, is then used as an indicator of anomaly. A high reconstruction error suggests that the data is anomalous, as it deviates from the learned normal patterns. Deep learning-based reconstructive models have become prominent due to their ability to capture complex patterns in high-dimensional data. In recent years, most reconstruction-based AD models have been developed using techniques such as GAN, AE, and diffusion models. These models each have unique strengths and weaknesses, as summarized in Table I. This table consolidates insights from multiple studies, including [80], [81], [82], and [83], which have analyzed the advantages and limitations of GANs, VAEs, and Diffusion Models in AD. In this section, we introduce these three types of models in the context of AD and discuss their various variants.

1) *GAN-based Anomaly Detection*: GANs are powerful tools for generating synthetic data that resembles a given training dataset [84]. As shown in the upper part of Fig.3, GANs consist of two main components: a generator and a discriminator, both of which are neural networks. Because of this structure, GAN models are highly flexible, allowing for different networks to be chosen as the generator and discriminator based on the specific task. This flexibility makes GANs a versatile framework for a wide range of applications. The generator  $G$  takes a random noise vector  $z$  (usually sampled from a Gaussian distribution) as input and generates synthetic data  $G(z)$ . The discriminator  $D$  receives a data sample (either from the real dataset or from the generator) as input and outputs a probability  $D(x)$ , representing the likelihood that the input is real (i.e., from the actual dataset) rather than fake (i.e.,

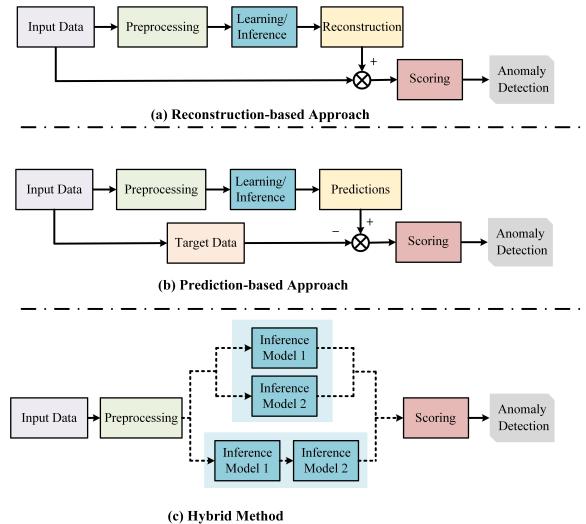


Fig. 2. Three types of anomaly detection: (a) Reconstruction-based approach, (b) Prediction-based approach, (c) Hybrid method.

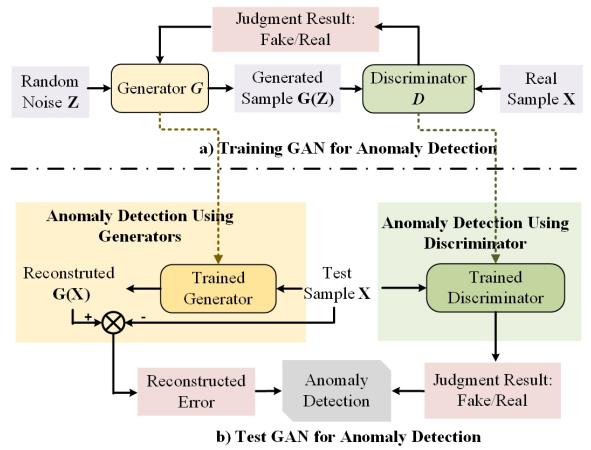


Fig. 3. Structural Frameworks for GAN Anomaly Detection.

generated by the generator). The generator and discriminator are trained simultaneously through a process where the generator tries to produce data that can fool the discriminator, and the discriminator tries to improve its ability to distinguish between real and fake data. Table II provides a comprehensive summary of recent GAN-based AD models, categorizing them based on their techniques, approaches, strengths, and weaknesses. This table highlights how different GAN variants are tailored for specific AD tasks, along with the types of data they are applied to and their publication years.

The training process of GANs can be described as a minimax game with the following objective function:

$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim p_{data}(x)}[\log D(x)] + \mathbb{E}_{z \sim p_z(z)}[\log(1 - D(G(z)))] \quad (1)$$

In this function,  $p_{data}(x)$  represents the distribution of the real data,  $p_z(z)$  represents the distribution of the noise vector  $z$ ,  $G(z)$  is the data generated by the generator, and  $D(x)$  is the

TABLE I  
COMPARISON OF GANs, VAEs, AND DIFFUSION MODELS IN ANOMALY DETECTION

Model	Strengths	Weaknesses
GANs	<ul style="list-style-type: none"> <li>Capable of generating high-fidelity, realistic samples.</li> <li>Learns complex data distributions using adversarial loss.</li> <li>Useful in AD by distinguishing real vs. generated data.</li> </ul>	<ul style="list-style-type: none"> <li>Prone to mode collapse, leading to low sample diversity.</li> <li>Hard to train with difficult-to-interpret losses.</li> <li>Training is unstable and hard to converge.</li> </ul>
VAEs	<ul style="list-style-type: none"> <li>Easy to train with one tractable likelihood loss.</li> <li>Provides high sample diversity by covering all data modes.</li> <li>Latent space representation is useful for AD tasks.</li> </ul>	<ul style="list-style-type: none"> <li>Produces low-fidelity, often blurry samples.</li> <li>Pixel-based loss leads to sample ambiguity and blurriness.</li> </ul>
Diffusion Models	<ul style="list-style-type: none"> <li>Generates high-fidelity samples with gradual refinement.</li> <li>High sample diversity due to likelihood maximization.</li> <li>Intermediate noisy images serve as useful latent codes for AD.</li> </ul>	<ul style="list-style-type: none"> <li>Slow sample generation due to the multi-step denoising process.</li> <li>Computationally intensive, requiring many steps for both forward and reverse diffusion.</li> </ul>

probability that  $x$  is real. The generator  $G$  aims to minimize this objective, while the discriminator  $D$  aims to maximize it. The discriminator updates its weights to maximize the probability of correctly classifying real and generated data, while the generator updates its weights to minimize the discriminator's ability to distinguish between real and fake data.

In the context of AD, GANs play crucial roles in both representation learning and data augmentation, each serving distinct purposes within deep Learning [85]. In representation learning, the primary objective of GANs is to learn and model the underlying distribution of the data, enabling the generation of synthetic data that closely resembles real data. This process involves a generator that creates fake data from random noise and a discriminator that distinguishes between real and fake data. Through iterative training, the generator improves its ability to produce realistic data, which is particularly useful in tasks like AD. For example, in [86], GANs are used for representation learning by generating fake data that matches the distribution of normal data. This generated data is then used to train a VAE to detect anomalies through reconstruction errors. Similarly, in [87], a fault-attention generative probabilistic adversarial autoencoder (FGPAA) is proposed, combining GANs and autoencoders for AD by learning the low-dimensional manifold of healthy state data. The GAN component aids in feature representation learning, reducing signal information loss and enhancing the model's ability to detect anomalies through distribution probability and reconstruction error.

There are two main structures to using GANs for AD, as shown in Fig.3. The first approach is based on the generator, as depicted in the lower part of Fig.3, highlighted by the yellow box. The basic idea is to train the GAN on normal data and then use the reconstruction error to identify anomalies. During the training phase, the GAN is trained exclusively on normal data, allowing the generator to learn to produce data that closely mimics the normal data distribution. During the detection phase, a test data point  $x$  is fed into the generator to obtain the reconstructed data  $G(x)$ . The reconstruction error, typically measured as the difference between the original data point  $x$  and the reconstructed data  $G(x)$ , is then used to detect anomalies. This can be quantified using metrics such as mean squared error (MSE). If the reconstruction error exceeds a predefined threshold, the data point is classified as an anomaly. The intuition behind this approach is that the generator, trained solely on normal data, will struggle to accurately reconstruct anomalous data, resulting in a high reconstruction error.

The mathematical representation for AD using GANs involves computing the reconstruction error  $E(x)$  as follows:

$$E(x) = \|x - G(x)\|^2, \quad (2)$$

where  $\|\cdot\|^2$  denotes the squared Euclidean distance. A threshold  $\tau$  is set, and if  $E(x) > \tau$ , the data point  $x$  is considered an anomaly. For example, Dong *et al.* [88] propose a semi-supervised approach for video AD using a dual discriminator-based GAN structure, focusing on representation learning. In this approach, the generator predicts future frames for normal events, and anomalies are detected by evaluating the quality of these predictions. Similarly, Guo *et al.* [89] introduce ReGraphGAN, a graph generative adversarial network specifically designed for dynamic graph AD. ReGraphGAN utilizes GAN-based representation learning to encode complex spatiotemporal relationships in graph data, allowing it to better capture anomalies. By leveraging encoders to project input samples into a latent space and integrating GANs to enhance both training stability and efficiency, ReGraphGAN significantly improves AD performance over existing methods.

The second approach leverages the discriminator highlighted by the green box in Fig.3. A well-trained discriminator has the ability to differentiate between real (normal) and fake (anomalous) samples. During the detection phase, test samples are directly input to the discriminator, which evaluates the likelihood that a given sample is real. If the discriminator assigns a low probability to a sample, suggesting that it is likely fake or anomalous, the sample is flagged as an anomaly. This method relies on the discriminator's capacity to recognize deviations from the normal data distribution it learned during training. For instance, Liu *et al.* [90] propose a GAN framework that uses multiple generators to produce potential outliers, which are then distinguished from normal data by a discriminator to detect anomalies. The discriminator's output score is used to evaluate the anomaly degree of input data, providing a comprehensive reference distribution and preventing mode collapse.

Additionally, GANs are highly effective in data augmentation, helping to mitigate the scarcity of anomaly samples, which often results in data imbalance and poor generalization [91]. When anomaly samples are unevenly distributed or lacking in diversity, models struggle to learn rare anomalies and can overfit to the training set, reducing their accuracy on unseen data. Traditional data augmentation techniques—such as scaling, rotation, random cropping, translation, flipping, and copy-paste—attempt to mitigate these issues. However, simple linear transformations fail to capture new distributions and features of unknown anomalies, such as random changes in

shape or texture. This is where GANs provide a significant advantage. By generating synthetic anomaly data that mimics the distribution of real-world anomalies, GANs enable models to learn a more diverse set of anomaly features. This not only addresses the imbalance problem but also improves the model's generalization capabilities, as it learns to detect anomalies based on a broader range of characteristics beyond those present in the original training dataset. Miao *et al.* [92] introduce an unsupervised AD framework that uses data augmentation through contrastive learning and GANs to mitigate overfitting. By employing a geometric distribution mask, it enhances data diversity and generates synthetic anomaly samples, addressing the scarcity of anomaly data. In [93], Anomaly-GAN addresses data augmentation by using a mask pool, anomaly-aware loss, and local-global discriminators to generate high-quality, realistic synthetic anomalies with diverse shapes, angles, spatial locations, and quantities in a controllable manner. Li *et al.* [94] propose augmented time regularized generative adversarial network that combines an augmented filter layer and a novel temporal distance metric to generate high-quality and diverse artificial data, addressing the limitations of existing GAN approaches in handling limited training data and temporal order.

2) *AE-based Anomaly Detection*: In recent years, the limitations of traditional AE models in handling complex and noisy data have become more apparent, leading to the development of enhanced methods to improve their performance in AD tasks. For example, Fan *et al.* [97] introduce a new framework by incorporating  $\ell_{2,1}$ -norm into the AE, and experiments have demonstrated that this framework can significantly improve ADn accuracy by increasing the model's robustness to noise and outliers during training. Wang *et al.* [98] demonstrate that introducing an adaptive-weighted loss function can effectively suppress anomaly reconstruction, thereby improving the accuracy of AD. Liu *et al.* [99] introduce a multi-scale convolutional AE architecture, where multiple stacked convolutional encoder-decoder layers act as background learners to robustly eliminate anomalies of varying sizes during background reconstruction. Additionally, Lin *et al.* [100] introduce a soft calibration strategy combined with AE to address the issue of data contamination in AD.

VAEs are another generative model widely used in AD tasks. Like GANs, VAEs aim to learn the distribution of normal data to identify anomalies. However, unlike GANs, which rely on adversarial training between a generator and a discriminator, VAEs use an encoder-decoder architecture. Fig.4 illustrates the structure of AD based on VAE. The goal of a VAE is to map the input data into a latent space through the encoder and model the data distribution probabilistically within this space. This approach allows the VAE to generate new data that closely resembles the true data distribution, and anomalies can be detected by evaluating the reconstruction error.

The internal structure of a VAE is similar to that of a traditional AE but with some key differences. First, the encoder in a VAE not only compresses the input data into a lower-dimensional latent space but also learns a probabilistic distribution, typically parameterized by a mean  $\mu$  and a vari-

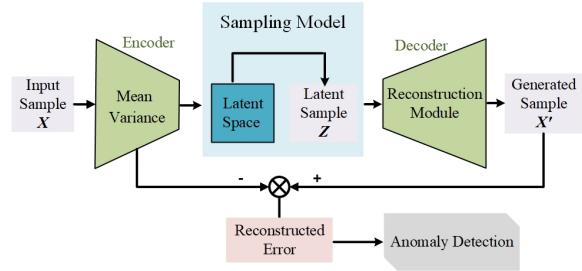


Fig. 4. Structural Frameworks for VAE Anomaly Detection.

ance  $\sigma^2$  as shown in Fig.4. This enables the VAE to generate more meaningful latent variables  $z$ , enhancing the diversity and robustness of the generated data. A critical component introduced in VAEs is the Kullback-Leibler (KL) divergence, which measures the difference between the latent distribution generated by the encoder and a predefined prior distribution (usually a standard normal distribution). Unlike traditional AEs, which focus solely on minimizing the reconstruction error, VAEs are trained by minimizing a combination of the reconstruction error and the KL divergence:

$$\mathcal{L}_{\text{VAE}} = \mathbb{E}_{q(z|x)}[\log p(x|z)] - D_{\text{KL}}(q(z|x)\|p(z)). \quad (3)$$

This difference makes VAEs more powerful in AD because they not only consider the quality of the data reconstruction but also enforce a structured latent space through the KL divergence. By doing so, KL divergence helps to regularize the latent space, ensuring that the encoded representations are smoothly distributed and centered around the prior distribution. This regularization reduces overfitting, promotes better generalization, and makes it easier to distinguish between normal and anomalous data, especially in complex and high-dimensional datasets. Table III provides a comprehensive summary of the latest advancements in VAE-based AD models, showcasing innovative enhancements that address various challenges such as noise robustness, semantic feature learning, and anomaly reconstruction. Huang *et al.* [101] enhance VAE-based AD by incorporating an Autoencoding Transformation into the model, which ensures that the training phase effectively captures high-level visual semantic features of normal images, thereby increasing the anomaly score gap between normal and anomalous samples. Similarly, Yin *et al.* [102] utilize Convolutional Neural Network (CNN) and VAE with a two-stage sliding window approach in data preprocessing to learn better representations for AD tasks. Zhang Yin *et al.* [103] propose the Graph Relational Learning Network (GReLeN), which integrates a VAE structure with graph dependency learning for AD in multivariate time series through reconstruction. Zhou *et al.* [104] propose a variational long short-term memory (VLSTM) model for high-dimensional AD in imbalanced datasets, combining a compression network for efficient data representation with an estimation network for accurate classification of network traffic data. The VLSTM model balances data compression and feature retention using core LSTM and variational modules.

In recent years, many advancements in AD models inspired by VAEs have focused on Adversarial Autoencoders (AAEs)

TABLE II  
GAN-BASED MODELS IN ANOMALY DETECTION

Paper	Technique	Approach Type	Strength	Weakness	Data Type	Year
[90]	GAN	Reconstruction	Does not depend on assumptions about the normal data and requires less computing resources.	The method involves the selection of multiple hyperparameters, making the tuning process challenging and potentially time-consuming.	Structured data	2020
[48]	GAN+CNN	Prediction	The NM-GAN model enhances both the generalization and discrimination abilities through noise-modulated adversarial learning, resulting in improved accuracy and stability for video AD.	The model struggles to fully capture complex temporal patterns like staying, wandering, and running, and lacks adaptive modulation of generalization and discrimination abilities, leaving room for improvement in spatiotemporal feature learning.	Video data	2021
[94]	GAN	Reconstruction	Is capable of generating more effective artificial samples for training supervised learning models, thereby addressing the issue of data imbalance.	Its performance is inferior to the baseline algorithms when the balanced ratio is 0.125.	Image data	2021
[95]	GAN+LSTM	Prediction	The TMANomaly framework excels in capturing complex multivariate correlations in industrial time series data, enhancing AD accuracy through mutual adversarial training.	The paper lacks discussion on TMANomaly's generalization to other datasets, the potential limitations of using GRA for feature selection, and the computational efficiency or scalability, which are critical for real-time industrial systems.	Multivariate time series data	2022
[96]	GAN+LSTM	Prediction	FGANomaly method effectively filters anomalous samples before training, improving AD accuracy and robustness by precisely capturing normal data distribution and dynamically adjusting generator focus.	The method lacks effective fusion of information across different dimensions in multivariate time series, which limits its ability to fully capture complex correlations.	Multivariate time series data	2022
[93]	GAN	Reconstruction	Improves the quality of the generated anomaly images and generates anomalies with different shapes, rotation angles, spatial locations, and numbers in a controllable manner.	The images generated are not very sensitive to the change of light.	Image data	2023
[89]	GAN	Reconstruction	Improves training efficiency and stability in dynamic graph AD while avoiding the expensive optimization process typical of traditional graph generative adversarial networks.	The detection accuracy on the UCI Message dataset is lower than that of TADDY.	Dynamic graph data	2023
[92]	GAN+Transformer	Reconstruction	It can effectively detect anomalies in long sequences, mitigates overfitting, and incorporates contrastive loss into the discriminator to fine-tune the GAN, ensuring strong generalization ability.	It may struggle with irregularly sampled data or datasets with many missing values, requires careful tuning of several hyperparameters, and demands significant computational resources, posing challenges for real-time processing on limited-capacity devices.	Multivariate time series data	2024

[105]. Unlike traditional VAEs, which use KL divergence to match the latent space distribution to a prior, AAEs achieve this through the use of GANs. Specifically, AAEs employ a GAN's discriminator to evaluate the latent variable distribution produced by the encoder and use adversarial training to align it with the desired prior distribution, providing more flexible control over the quality of the generated data. Wu *et al.* [87] propose the Fault-Attention Generative Probabilistic Adversarial Autoencoder (FGPAA) for machine AD, utilizing an end-to-end AAE with double discriminators to extract relevant features and ensure accurate equipment health monitoring through a fault-attention probability distribution. Idrissi *et al.* [51] apply AAE and FL in the field of network intrusion detection, effectively ensuring AD performance while safeguarding client privacy. Experimental results demonstrate that the proposed model outperforms AE, VAE, and AAE on various network traffic datasets, achieving high performance across different metrics. Su *et al.* [106] propose two contamination-immune BiGAN models, integrating elements of VAE and BiGAN to create a new AAE-based framework that effectively detects anomalies by learning the probability distribution of normal samples from contaminated datasets, significantly outperform-

ing state-of-the-art methods in scenarios where training data is impure. Similar to the aforementioned AAE models, Du *et al.* use GANs to purify the original dataset, generating synthetic "normal" data to improve outlier detection accuracy. Continuing the advancements in AAE-based models, Yu *et al.* [107] introduce an Adversarial Contrastive Autoencoder (ACAE) for Multivariate Time Series (MTS) AD, which enhances feature representation through adversarial training and contrastive learning, demonstrating superior performance across multiple real-world datasets, further extending the application of AAE-based methods in robust AD.

3) *Diffusion model-Based for Anomaly Detection:* Diffusion models are a type of generative model that operate through two key phases: a fixed forward diffusion process and a learnable reverse diffusion process [108]. Mathematically, the forward process involves progressively adding Gaussian noise to the data  $x_0$ , transforming it into pure noise  $x_T$  over  $T$  steps. This process can be described as:

$$q(x_t|x_{t-1}) = \mathcal{N}(x_t; \sqrt{1 - \beta_t}x_{t-1}, \beta_t I), \quad (4)$$

where  $q(x_t|x_{t-1})$  is the conditional probability distribution of  $x_t$  given  $x_{t-1}$ ,  $\beta_t$  is the noise variance at step  $t$ , and  $x_t$

TABLE III  
AUTOENCODER-BASED MODELS IN ANOMALY DETECTION

Paper	Technique	Approach Type	Strength	Weakness	Data Type	Year
[104]	VAE-based (VAE+LSTM)	Reconstruction	Effectively addresses imbalanced and high-dimensional challenges in industrial big data.	Falls short in achieving the highest AUC and F1 scores compared to other methods.	Industrial big data	2020
[87]	AAE-based	Reconstruction	FGPAA reduces information loss during feature extraction and constructs fault attention anomaly indicators using low-dimensional feature probability and reconstruction error.	Runtime is approximately five times longer than SOM.	Rotating machine fault simulator data	2020
[98]	AE-based (AE+CNN)	Reconstruction	The Auto-AD method enables fully autonomous hyperspectral AD, automatically separating anomalies based on reconstruction errors without the need for manual tuning or additional processing.	Lower AUC score compared to the GRX method on the Honghu dataset.	Hyperspectral data	2021
[99]	AE-based (AE+CNN)	Reconstruction	MSNet offers an effective solution to handle multiscale anomaly shapes, providing greater flexibility without the need for threshold fine-tuning.	Multiple convolutional encoder-decoder layers and enhanced training increase computational cost and training time.	Hyperspectral data	2021
[101]	VAE-based (VAE+Transformer)	Reconstruction	SSR-AE leverages self-supervised learning to enhance normal data reconstruction and hinder abnormal data, optimizing mutual information for effective transformation and image reconstruction.	Struggles with transformations, heavily relying on their effectiveness for AD.	Image data	2021
[97]	AE-based	Reconstruction	Maintains geometric structure and local spatial coherence of hyperspectral images (HSI), reducing search space and execution time per pixel.	High execution time for constructing the SuperGraph matrix with large datasets.	Hyperspectral data	2021
[51]	AAE-based (AAE+Federated learning)	Reconstruction	Fed-ANIDS demonstrates strong generalization, outperforms GAN-based models, and ensures privacy protection through federated learning.	Computational overhead due to the federated learning framework, increasing training complexity and latency.	Cybersecurity data	2023
[100]	AE-based	Reconstruction	Applicable for time series AD under data contamination.	Assumes normal samples follow a Gaussian distribution, limiting applicability, and has higher computational complexity.	Time series data	2024
[106]	AAE-based	Reconstruction	Learns the probability distribution of normal samples from contaminated datasets, achieving convergence and outperforming baseline models.	Relies on the assumption that the contamination ratio is known, which may not always be accurate in practice.	Medical image data	2024
[86]	AAE-based	Reconstruction	Generates a clean dataset from contaminated data for AD, with linear scalability for larger datasets.	Struggles with detection accuracy in datasets with multiple distribution patterns.	Tabular data	2024
[107]	AAE-based	Reconstruction	Excels in learning high-level semantic features and capturing normal patterns of MTS with contrastive learning constraints, ensuring stability across parameter settings.	Performance on all metrics for SMAP and PSM datasets is lower than baseline methods.	Multivariate time series data	2024

represents the noisy data at step  $t$ . As  $t$  increases, the data becomes more corrupted by noise until it reaches a state of pure Gaussian noise at step  $T$ .

The reverse process learns to gradually denoise the data, removing the added noise step by step. The model learns a parameterized distribution  $p_\theta(x_{t-1}|x_t)$  to reverse the noise addition process, reconstructing the original data from the noisy data. This reverse process is trained to minimize the variational bound on the data likelihood, expressed as:

$$L = \mathbb{E}_q [D_{KL}(q(x_{t-1}|x_t, x_0) || p_\theta(x_{t-1}|x_t))]. \quad (5)$$

By progressively removing noise, diffusion models generate high-fidelity samples, first capturing coarse structures and then refining details in each step. In the context of AD, diffusion models are trained on normal data to learn the underlying data distribution through an iterative noise-removal process. Similar to other reconstruction-based methods, anomalies can be identified by evaluating the reconstruction error, where a higher error indicates that the data deviates from the learned normal patterns.

Diffusion models stand out from GANs and VAEs in several key ways. They avoid common issues such as mode collapse in GANs, where only a subset of the data distribution is captured, leading to reduced diversity. Diffusion models also overcome the blurriness associated with VAEs, which often results from pixel-based loss and a smaller latent space. By iteratively denoising data, diffusion models maintain both high fidelity and diversity in their outputs.

While diffusion models are slower in generating samples due to their iterative nature, their ability to accurately reconstruct data and cover the full range of the training dataset makes them particularly well-suited for AD [109]. In AD, where precision is critical, diffusion models excel by generating detailed and high-quality samples, enabling them to identify subtle deviations from normal patterns with greater accuracy than other generative models. Several works have leveraged the advantages of diffusion models in ADn. For example, Zhang *et al.* [110] utilize the high-quality and diverse image generation capabilities of diffusion models to enhance reconstruction quality in DiffAD, addressing the limitations of

traditional methods by introducing noisy condition embedding and interpolated channels. Similarly, Li *et al.* [111] apply a diffusion model to reconstruct normal data distributions and integrate an auxiliary learning module with pretext tasks to better distinguish between normal and abnormal data. Expanding on these ideas, Zeng *et al.* [112] improve denoising diffusion probabilistic models (DDPMs) for radio AD by incorporating an AE to learn the distribution of normal signals and their power spectral density (PSD), using reconstruction error to identify anomalies. Li *et al.* [113] present a Controlled Graph Neural Network (ConGNN) approach based on DDPMs to address the challenge of limited labeled data. Li *et al.* [114] further explore diffusion models in vehicle trajectory AD, employing decoupled Transformer-based encoders to capture temporal dependencies and spatial interactions among vehicles, significantly improving AUC and F1 scores on real-world and synthetic datasets. Similarly, Pei *et al.* [115] establish the two-stage diffusion model (TSDM) to mitigate the influences of anomalies in smart grids, where the first stage is a diffusion-based AD component. In multi-class AD, He *et al.* [116] propose DiAD, a framework that enhances reconstruction accuracy through a combination of a semantic-guided network, spatial-aware feature fusion, and a pre-trained feature extractor to generate anomaly maps.

#### B. Deep learning methods for Anomaly Detection based on Prediction

Prediction-based AD methods operate by forecasting future values or estimating missing attributes and comparing these predictions to the actual observed values. When significant deviations occur, it indicates potential anomalies, as the data deviates from the learned normal patterns. These methods are versatile and can be applied across various data types, leveraging relationships between variables or temporal correlations to detect anomalies. Prediction-based methods excel in scenarios where capturing patterns and trends is essential. By learning underlying structures in the data, whether based on time dependencies or more general interactions between variables, these methods can effectively predict expected outcomes. Deviations from these expectations are flagged as anomalies. This makes prediction-based approaches highly adaptable, capable of functioning across different contexts, including various types of data. In this section, we explore three main approaches for prediction-based AD: Recurrent Neural Networks (RNNs), attention mechanisms, and Graph Neural Networks (GNNs), all of which have demonstrated efficacy in capturing intricate patterns and relationships within data to identify anomalies. These methods allow for flexible and robust AD across various data types by learning underlying patterns, whether they are based on spatial, temporal, or graph-based relationships. By leveraging these approaches, prediction-based methods can effectively model complex interactions, providing reliable detection of unexpected behaviors or deviations from learned patterns.

1) *RNN-based Anomaly Detection:* Recurrent Neural Networks (RNNs) [117] are a special type of neural network designed to process sequential data by capturing dependencies between elements in a sequence. Unlike standard neural

networks, RNNs incorporate a state vector  $s_t$  in the hidden layer, allowing them to retain information from previous steps and model sequential patterns. This capability makes them effective in various applications where data has an inherent order, such as event logs, system monitoring, and structured sequences in cybersecurity or industrial processes. For an input  $x_t$  at time  $t$ , the update of the state value  $s_t$  and hidden layer output  $h_t$  in RNNs can be represented as

$$\begin{aligned} s_t &= \sigma(\mathbf{W}^x x_t + \mathbf{W}^s s_{t-1} + \mathbf{b}^s) \\ h_t &= \text{softmax}(\mathbf{W}^h s_t + \mathbf{b}^h), \end{aligned} \quad (6)$$

where  $\sigma(\cdot)$  is the sigmoid activation function,  $\mathbf{W}^x$ ,  $\mathbf{W}^s$  and  $\mathbf{W}^h$  represent the network weights, and  $\mathbf{b}$  is the network biases. By maintaining a recurrent state, RNNs can effectively capture dependencies across different steps within a sequence, making them well-suited for tasks involving ordered data.

However, RNNs face the problem of exploding or vanishing gradients when dealing with long sequences. Long Short-Term Memory networks (LSTMs) [118], a specialized type of RNN, were introduced to address these issues. Specifically, LSTMs replace the hidden layer of RNNs with an LSTM block consisting of input, output, and forget gates. The inference process of LSTM at time  $t$  is given by

$$\begin{aligned} f_t &= \sigma(\mathbf{W}^{xf} x_t + \mathbf{W}^{hf} h_{t-1} + \mathbf{b}^f) \\ i_t &= \sigma(\mathbf{W}^{xi} x_t + \mathbf{W}^{hi} h_{t-1} + \mathbf{b}^i) \\ \tilde{c}_t &= \tanh(\mathbf{W}^{x\tilde{c}} x_t + \mathbf{W}^{h\tilde{c}} h_{t-1} + \mathbf{b}^{\tilde{c}}) \\ c_t &= f_t c_{t-1} + i_t \tilde{c}_t \\ o_t &= \sigma(\mathbf{W}^{xo} x_t + \mathbf{W}^{ho} h_{t-1} + \mathbf{b}^o) \\ h_t &= o_t \tanh(c_t), \end{aligned} \quad (7)$$

where  $f_t$ ,  $i_t$ , and  $o_t$  are the forget, input and output gate weights, respectively.  $c_t$  represents the cell state of LSTM, and  $\tanh(\cdot)$  is the hyperbolic tangent activation function. By controlling the weights of the forget, input, and output gates, LSTM determines the importance of historical time series information and the current input on the current output, thus effectively mitigating issues of gradient vanishing and allowing robust modeling of complex sequences. Reference [119] provides comprehensive evidence of LSTM's effectiveness in AD across various technical systems, demonstrating its superiority in learning complex temporal behaviors and accurately identifying anomalies.

The Gated Recurrent Unit (GRU) [120] is a simplified version of LSTM that only includes an update gate and a reset gate and uses the hidden state alone to represent both short-term and long-term information. These different types of RNNs can be used in prediction-based AD tasks, with the specific detection and inference method illustrated in Fig. 5. RNNs, LSTMs, and GRUs take time series data from  $t-w$  to  $t-1$  as input, and their pre-trained neural networks use these temporally ordered data to predict the single-step or multi-step future values of the univariate or multivariate time series. If the difference between the actual and predicted values is below a threshold, no anomaly is detected; if the difference exceeds the threshold, an anomaly is detected and the spatiotemporal location of the anomaly is identified.

TABLE IV  
DIFFUSION-BASED MODELS IN ANOMALY DETECTION

Paper	Technique	Approach Type	Strength	Weakness	Data Type	Year
[110]	Diffusion	Reconstruction	The latent diffusion model (LDM) used in this method achieves state-of-the-art performance in surface AD by generating high-quality, semantically correct reconstructions, effectively avoiding overfitting to anomalies.	It is less suitable for real-time applications or environments with limited computational resources.	Image data	2023
[112]	Diffusion+VAE	Reconstruction	The AE-DDPMs algorithm effectively improves stability and reduces computational costs in radio AD, outperforming GAN-based methods in complex electromagnetic environments.	The anomalies in the experimental data are artificially generated, rather than originating from real-world conditions, which may limit the model's applicability to genuine, real-world scenarios.	radio signal data	2023
[113]	Diffusion+GNN	Prediction	ConGNN effectively addresses the issue of limited labeled data by generating augmented graph data using a graph-specific diffusion model.	The reliance on graph-specific augmentation might not generalize well to other types of data, potentially limiting its applicability beyond graph-based AD.	Image data	2023
[111]	Diffusion+VAE	Hybrid	SDAD effectively enhances AD by combining self-supervised learning for discriminative data representation with denoising diffusion.	The generation of pseudo anomalies relies solely on standard Gaussian sampling, which may not fully capture the complexity of real anomalies, limiting the model's ability to accurately simulate genuine abnormal data.	Structure data	2024
[114]	Diffusion+Transformer	Hybrid	DiffTAD effectively models temporal dependencies and spatial interactions in vehicle trajectories through diffusion models, significantly improving AD accuracy and robustness to noise.	The anomalies are primarily evaluated on synthetic datasets, which may not fully reflect the complexity and diversity of real-world trajectory data.	Vehicle trajectory data	2024

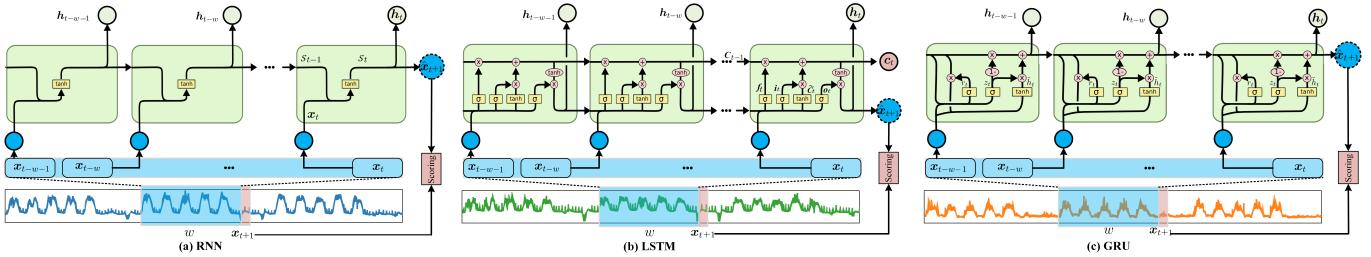


Fig. 5. RNN-based application example for time series data anomaly detection: (a) RNN-based, (b) LSTM-based, (c) GRU-based.

Current RNN-based AD primarily focuses on improving RNN algorithms tailored to AD tasks and integrating RNN with other methods for AD. The method in [121] employs a pruning algorithm to reduce the number of false data points, enabling the LSTM-based AD approach to better address the challenges posed by the extremely uneven distribution of railway traffic data. LSTM combined with AE [122], VAE [123], and Singular Value Decomposition (SVD) [124] has also been used to identify anomalies in Controller Area Networks (CANs) [125], electrocardiograms, and Internet monitoring data. GANs based on adversarial learning have also been integrated into the time series learning of LSTM, achieving very high performance in scenarios with few features [95], extremely imbalanced training sets, and noise interference [96]. CNN is also integrated into LSTM in a serial [126], parallel [127], or as a foundational layer [128] to better extract the spatiotemporal correlations of multidimensional time series, thereby enhancing the performance of AD. GRUs, compared to LSTMs, have a more streamlined architecture, resulting in lower computational complexity during training and execution of AD tasks, and they tend to perform better on certain less complex sequential data. For instance, GRUs enhance interpretability by uncovering latent correlations in

multivariate time series data from industrial control system sensors [129]. Similar to LSTMs, GRUs can also be combined with AEs [130] or VAEs [25] in an encoder-decoder architecture to mitigate the effects of noise and anomalies, thereby improving the accuracy of AD.

2) *Attention-based Anomaly Detection*: The attention mechanism was initially applied in machine translation [131], with its core idea being to enable the neural network to focus on the relevant parts of the input values. While attention-based methods have shown great promise in time series AD, their applications are not limited to temporal data. These methods can effectively capture dependencies in various types of data, including spatial, spatiotemporal, and multimodal datasets. This flexibility broadens their use cases across different AD tasks. Compared to RNN-based approaches, they are better suited for long or complex sequences because attention can compute dependencies between all positions in the sequence simultaneously, while RNNs process sequences sequentially, step by step.

Figure 6 illustrates a typical attention-based model for AD. Among attention-based methods, the self-attention mechanism is particularly effective in capturing global dependencies across various types of sequential data, including temporal,

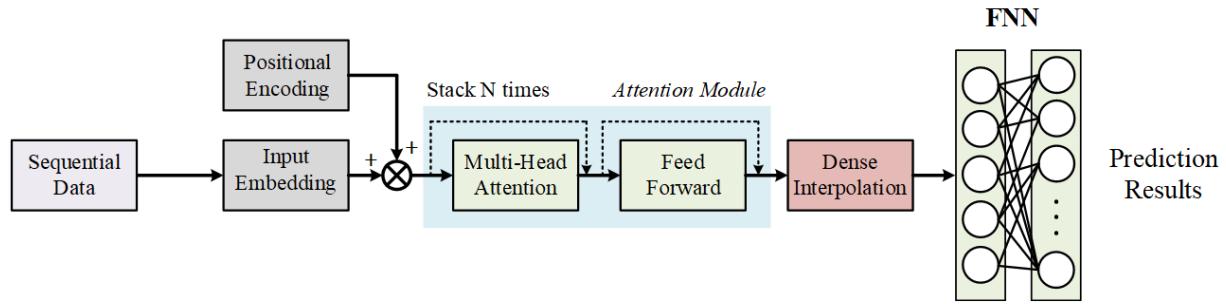


Fig. 6. Attention-based model for anomaly detection. The model first embeds sequential data using input embedding and positional encoding to preserve temporal dependencies. The multi-head attention mechanism captures long-range dependencies by processing interactions between all time steps. The feed-forward layer then refines feature representations, and a dense interpolation layer enhances anomaly-related features before passing them to a fully connected network (FNN) for final AD.

spatial, and spatiotemporal inputs. For an input dataset  $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_t]$ , the queries, keys, and values are defined as:  $\mathbf{Q} = \mathbf{X}\mathbf{W}_Q$ ,  $\mathbf{K} = \mathbf{X}\mathbf{W}_K$ , and  $\mathbf{V} = \mathbf{X}\mathbf{W}_V$ , where  $\mathbf{W}_Q$ ,  $\mathbf{W}_K$ , and  $\mathbf{W}_V$  are trainable weight matrices. The attention weights are then computed based on  $\mathbf{Q}$ ,  $\mathbf{K}$ , and  $\mathbf{V}$  as

$$\alpha_{ij} = \frac{\exp(\mathbf{Q}_i \mathbf{K}_j^\top / \sqrt{d_k})}{\sum_{j=1}^T \exp(\mathbf{Q}_i \mathbf{K}_j^\top / \sqrt{d_k})}, \quad (8)$$

where  $d_k$  is the dimension of the keys. Finally, the output of the self-attention-based neural network, which takes into account the importance of each input value, is given by  $\text{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \alpha\mathbf{V}$ .

To enable the model to capture features of various patterns, multi-head attention is also well-suited for AD. The calculation of multiple heads is expressed as

$$\text{Multihead}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{Concat}(\text{head}_1, \dots, \text{head}_h) \mathbf{W}_O, \quad (9)$$

where each head is computed as  $\text{head}_i = \text{Attention}(\mathbf{Q}\mathbf{W}_{Qi}, \mathbf{K}\mathbf{W}_{Ki}, \mathbf{V}\mathbf{W}_{Vi})$ . Here,  $\mathbf{W}_{Qi}$ ,  $\mathbf{W}_{Ki}$ , and  $\mathbf{W}_{Vi}$  are trainable parameters for different heads, and  $\mathbf{W}_O$  is the linear transformation matrix for the output.  $\text{Concat}(\text{head}_1, \dots, \text{head}_h)$  concatenates the outputs of all attention heads along the feature dimension. Attention-based methods can effectively capture long-term dependencies, improve computational efficiency, and enhance the interpretability of AD through visualized attention weight values. When applied to AD, differences in the distribution of attention weights between normal and anomalous time series can serve as the basis for AD.

In the field of AD, particularly for time series data, there has been a growing number of studies proposing deep learning methods based on attention mechanisms. Autoencoders that combine convolution, LSTM, and self-attention mechanisms can better extract complex features from multivariate time series data and robustly detect anomalies in high noise conditions [132]. The Transformer, as a well-known attention-based model, has demonstrated superior performance in unsupervised prediction-based time series AD compared to LSTM, as it can learn the dynamic patterns of sequential data through self-attention mechanisms [133]. The Transformer-based AD utilizes attention-based sequence encoders for rapid inference, achieving an F1 score improvement of up to 17%

on public datasets and reducing training time by as much as 99% compared to the baseline [134]. Despite its outstanding capabilities, the Transformer still faces certain bottlenecks in AD. Attention-based methods are prone to overfitting when data is insufficient. The method in [92] seamlessly integrates contrastive learning and GAN into the Transformer, utilizing data augmentation techniques and geometric distribution masking to expand the training data, thereby enhancing data diversity and improving accuracy by 9.28%.

Attention mechanisms are also frequently applied in graph neural networks to jointly detect anomalies in time series data. Reference [135] proposes a novel efficient Transformer model based on graph learning methods, employing two-stage adversarial training to train the AD model and utilizing prototypical networks to apply the model to anomaly classification. A contrastive time-frequency reconstruction network for unsupervised AD is used for AD and localization [136], where attention mechanisms and graph convolutional networks update the feature information of each time point, combining points with similar feature relationships to dilute the influence of anomalous points on normal points. Reference [137] models the correlations between temporal variables using graph convolutional networks, while also using an attention-based reconstruction model to output the importance of time series data within each time window, achieving an average AD F1 score exceeding 0.96. For multimodal data, a multimodal graph attention network (M-GAT) and temporal convolutional networks are used to capture spatial-temporal correlations in multimodal time series and correlations between modalities [138], ultimately outputting anomaly scores through reconstruction or prediction. More details about the application of GNNs in AD will be elaborated in the next subsection.

In addition to GNNs, CNNs can also incorporate attention mechanisms to enhance various metrics of AD. Reference [139] effectively captures the local features of subsequences by leveraging the locality of CNNs and combining it with positional embeddings. At the same time, Zhu *et al.* [139] employ attention mechanisms to extract global features from the entire time series, thereby enhancing the effectiveness and potential of detection. Many works have also introduced LSTM to extract temporal correlations in time series data based on CNN models with attention mechanisms. For example, Sun *et al.* [140] employ a sequential approach where 1D convolution is

first used to extract abstract features of the signal values at each time step, which are then input into a bidirectional long short-term memory network (Bi-LSTM), ultimately combining with attention mechanisms to make the model focus on locally important time steps. Meanwhile, Le *et al.* [141] integrate convolutional layers, LSTM layers, and self-attention layers into an autoencoder architecture to better extract complex features from multivariate time series. Similarly, Pei *et al.* [126] employ additional SVM to classify the attention weights based on a CNN-LSTM model with attention mechanisms to determine whether cyber-attacks have occurred in energy systems. The input data are the multimodal measurements from the deployed sensors.

3) *GNN-based Anomaly Detection:* Graph Neural Networks (GNNs) have gained increasing attention in AD tasks, as many types of data can be naturally represented as graph structures [142]. Wu *et al.* [143] have demonstrated the effectiveness of GNNs in identifying anomalies within complex graph-structured data environments. As neural network models specifically designed to handle graph-structured data, GNNs define nodes, edges, and graphs, where nodes represent individual elements in the dataset, such as data points in a sequence, sensor readings in multivariate data, or entities in relational datasets—denoted as the set  $V$ . Edges capture the relationships or dependencies between these elements, denoted as the set  $E$ , and can represent temporal correlations, spatial dependencies, or more abstract relational connections depending on the context. The graph, represented as  $G = (V, E)$ , captures the overall structure formed by nodes and edges. The primary operations in GNN training are message passing and aggregation, which are used to update and learn node features. Specifically, during message passing, each node receives information from its neighboring nodes and updates its own state. For a node  $v$ , the message passing formula is given as

$$\mathbf{m}_v^{(k)} = \sum_{u \in \mathcal{N}(v)} MSG\left(\mathbf{h}_u^{(k-1)}, \mathbf{h}_v^{(k-1)}, \mathbf{e}_{uv}\right), \quad (10)$$

where  $\mathcal{N}(v)$  denotes the set of neighboring nodes of  $v$ ,  $\mathbf{h}_u$  and  $\mathbf{h}_v$  are the features of nodes  $u$  and  $v$  at layer  $k$ , and  $\mathbf{e}_{uv}$  represents the edge features. Subsequently, the received messages are aggregated with the current node state, and the node features are updated as

$$\mathbf{h}_v^{(k)} = UPDATE\left(\mathbf{h}_v^{(k-1)}, \mathbf{m}_v^{(k)}\right), \quad (11)$$

where  $UPDATE(\cdot, \cdot)$  is the update function.

As illustrated in Fig. 7, which uses time series data as an example, GNNs treat each variable in the multivariate time series as a node to capture complex relationships between different dimensions. While the primary focus here is on the predictive capabilities of GNNs, it is worth noting that they are also effective in reconstruction-based AD. The final decision on whether the input sequence is anomalous is primarily based on prediction errors or graph structure differences, with reconstruction errors serving as a supplementary indicator. GNN-based AD methods excel at modeling complex dependencies between time steps or sensors, offering flexibility to handle

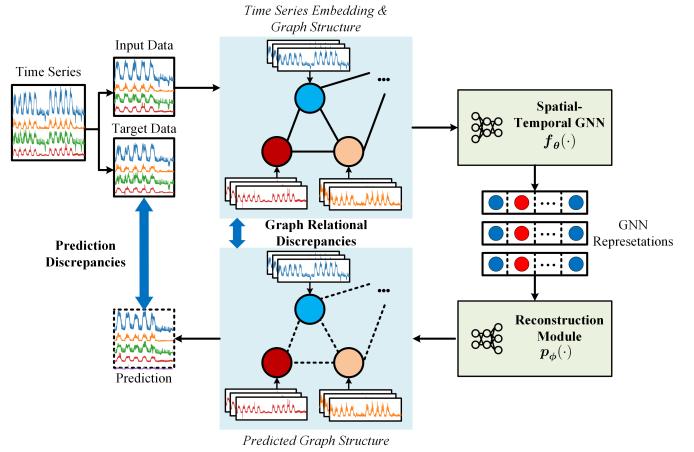


Fig. 7. GNN-based method for anomaly detection with time series data. Time series data is embedded into a graph structure, where a spatial-temporal GNN extracts dependencies. The reconstruction module then estimates the original data. Anomalies are detected based on graph relational discrepancies (differences in predicted graph structure) and prediction discrepancies (differences between reconstructed and actual time series).

both static and dynamic relationships across diverse time series structures. However, they still face challenges such as high computational complexity on large-scale graphs and difficulties in constructing optimal edge and graph configurations [144].

In prediction-based GNN for AD, GDN [145] is a representative work that combines a structure learning approach with GNN, additionally using attention weights to predict time series values and detect anomalies based on the predictions. Similar methods include GTA [146] and CST-GL [147]. Furthermore, Liu *et al.* [148] propose a GNN-based contrastive learning model that generates prediction scores from high-dimensional attributes and local structures to detect anomalies, outperforming state-of-the-art methods on seven benchmark datasets. Beyond prediction-based methods, there are also reconstruction-based GNN approaches. For example, MTAD-GAT [149] employs a graph attention network as a spatiotemporal encoder to learn dependencies across variables and time, reconstructing the time series with a backbone reconstructor and identifying anomalies based on reconstruction errors. Similar techniques include VGCRN [150] and FuSAGNet [151].

### C. Deep learning methods for Anomaly Detection based on Hybrid Method

In AD, reconstruction-based and prediction-based methods offer distinct but complementary approaches to identifying anomalies. Both methods rely on the discrepancy between the model's output and the actual input data as an indicator of abnormality. However, they diverge in how they handle data and their areas of application. Reconstruction-based methods focus on learning the underlying distribution of normal data. Once trained, the model attempts to recreate the input data. The reconstruction error, measured as the difference between the original data and its reconstruction, serves as a key indicator of anomalies. A high reconstruction error suggests that the data deviates from the normal patterns learned by the

model. This approach is particularly effective in cases where understanding the full structure or distribution of the data is crucial, such as in image-based AD or other high-dimensional datasets. In contrast, prediction-based methods focus on forecasting specific attributes or missing values from the data, rather than reconstructing the entire input. These methods typically predict future values or infer missing data points by leveraging known features. If the predicted values significantly deviate from the actual values, this signals a potential anomaly. Prediction-based methods are often more suited to feature-rich datasets, where predicting specific variables can help identify irregular patterns. For instance, in applications like fraud detection, predicting expected behaviors or transactions can reveal anomalies when the predicted outcomes differ from the observed ones. While both methods differ in their data processing approaches, they can be highly complementary. In many cases, combining reconstruction-based and prediction-based techniques within a hybrid framework allows for more robust AD. Reconstruction models capture the overall structure and patterns in the data, while prediction models focus on detecting deviations in specific variables or features. This combination can provide a more comprehensive solution for identifying anomalies in complex datasets across various domains. Tang *et al.* [152] utilize a U-Net module as the prediction module to perform future frame prediction, amplifying reconstruction errors for abnormal events, while another U-Net module is used as the reconstruction module to enhance predicted frames for normal events, thus improving the effectiveness of AD. Lv *et al.* [31] adopt a dilated convolution-based autoencoder to integrate prediction errors and reconstruction errors into the output anomaly scores, effectively improving the generalization capability of the detection model. Liu *et al.* [153] leverage a reconstruction model and a prediction model within an end-to-end semi-supervised AD framework to effectively capture inter-variable correlations and temporal dependencies in multivariate time series data from wind turbines. Additionally, by incorporating an auxiliary discriminator with adversarial training, the model can progressively improve performance using limited labeled data, enhancing the transition from unsupervised to supervised AD. Wei *et al.* [154] propose a hybrid deep-learning model combining LSTM and autoencoder for AD in indoor air quality data, where the LSTM captures long-term dependencies in time-series data and the autoencoder uses reconstruction loss to detect anomalies, effectively addressing both temporal correlations and reconstruction errors for improved detection accuracy.

#### D. Summary and Insights

This section introduces three types of deep learning-based AD methods: reconstruction-based, prediction-based, and hybrid approaches. Reconstruction-based methods are particularly effective in handling high-dimensional and unsupervised data by learning intrinsic patterns and identifying deviations through reconstruction errors. Prediction-based methods excel at modeling temporal dependencies in time-series data, enabling the detection of unexpected patterns in dynamic environments. Hybrid approaches combine these strengths

to address complex scenarios where multiple anomaly types coexist. Notably, these methods demonstrate the power of deep learning in capturing intricate patterns and dependencies that traditional methods often miss, making them indispensable for tackling diverse and challenging AD tasks.

## V. INTEGRATE TRADITIONAL METHOD AND DEEP LEARNING METHOD

In the field of AD, traditional methods and deep learning approaches each offer unique advantages. Traditional methods, such as clustering [155] and Support Vector Data Description [156], are often simpler, more interpretable, and computationally efficient. These methods excel in providing transparent decision-making processes, making them suitable for applications where model interpretability is crucial. On the other hand, deep learning methods, with their ability to model complex, high-dimensional data distributions, offer enhanced detection accuracy and adaptability, especially for large datasets and unstructured data like images and sequences.

The integration of traditional and deep learning methods aims to leverage the interpretability and simplicity of traditional methods with the robustness and flexibility of deep learning techniques. By combining these approaches, researchers seek to create hybrid models that maintain accuracy while offering insights into the underlying decision-making process, improving both detection power and model transparency.

### A. Clustering method

Clustering models play a crucial role in unsupervised AD, particularly for textual data. These models group similar data points based on their proximity in feature space and identify anomalies as points that deviate from established clusters [157]. Common clustering techniques, such as k-means [158], Density-Based Spatial Clustering of Applications with Noise (DBSCAN) [159], and hierarchical clustering [160], work effectively for simpler datasets and offer the advantage of interpretability. By integrating clustering methods with deep learning, such as applying clustering post feature extraction by a neural network, it is possible to improve detection accuracy while maintaining an interpretable clustering structure. This hybrid approach is particularly useful in cases where data distribution varies, and flexible, context-aware AD is required. For instance, Li *et al.* [161] propose a method that extends fuzzy clustering with a reconstruction criterion and Particle Swarm Optimization (PSO) to detect anomalies in both amplitude and shape. This highlights how traditional clustering methods can benefit from optimization techniques to handle diverse anomaly types. Similarly, Markovitz *et al.* [162] introduce an innovative approach for AD in human actions by working directly on human pose graphs extracted from video sequences. By mapping these graphs to a latent space, clustering them, and applying a Dirichlet process-based mixture model, the method effectively leverages probabilistic modeling to enhance the robustness and flexibility of clustering for action recognition. In video AD, Qiu *et al.* [163] propose a convolution-enhanced self-attentive video auto-encoder

integrated with a dual-scale clustering module based on the K-means algorithm. This approach effectively distinguishes normal and abnormal video data by enhancing feature representations and addressing the fuzzy boundaries between them. Additionally, Peng *et al.* [33] introduce a multivariate ELM-MI framework combined with a dynamic kernel selection method. By employing hierarchical clustering on unlabeled data to determine kernels, this method enables unsupervised online detection of various anomaly types, including point and group anomalies, while reducing computational costs and improving robustness. These studies collectively highlight the potential of hybrid approaches that integrate clustering with advanced techniques like deep learning, probabilistic modeling, or optimization frameworks. Such methods leverage the interpretability and simplicity of traditional clustering while addressing its limitations in handling complex data, offering a promising pathway for accurate and flexible AD.

### B. Normalizing Flows

Normalizing Flows (NF) [164] offer a probabilistic framework for AD by estimating the probability distribution of data. Using a sequence of invertible transformations, NFs can model complex distributions, making them particularly effective for identifying anomalies as low-probability events. When integrated with deep learning models, such as CNNs or RNNs, NFs act as precise probabilistic estimators, complementing the feature extraction capabilities of deep networks. This hybrid framework enhances AD, particularly in high-dimensional or unstructured datasets.

For instance, Yu *et al.* [165] propose FastFlow, a 2D normalizing flow module integrated with deep feature extractors like ResNet and Vision Transformers. By effectively modeling feature distributions and capturing both local and global relationships, FastFlow achieves state-of-the-art performance, with a 99.4% AUC on the MVTec AD dataset, while maintaining high inference efficiency. Similarly, Cho *et al.* [166] introduce Implicit Two-path Autoencoder (ITAE), which reconstructs normal video patterns by implicitly modeling appearance and motion features through two encoders and a shared decoder. NF enhances ITAE by estimating the density of normal embeddings, enabling robust detection of out-of-distribution anomalies, with strong results across six surveillance benchmarks. For multivariate time series data, Zhou *et al.* [167] combine a graph structure learning model with entity-aware normalizing flows to capture interdependencies and evolving relations among entities. By estimating entity-specific densities and employing a clustering strategy for similar entities, the extended MTGFlow\_cluster improves density estimation accuracy, demonstrating superior performance on six benchmark datasets. Further expanding on the use of graphs, Dai *et al.* [168] propose Graph-Augmented Normalizing Flow (GANF), which incorporates a Bayesian network to model causal relationships among time series. This approach factorizes joint probabilities into conditional probabilities, improving density estimation and enabling effective detection of anomalies in low-density regions, as well as identifying distribution drifts.

These studies collectively highlight the strengths of integrating Normalizing Flows with traditional and deep learning

based methods. By combining the interpretability and precision of probabilistic models with the expressive power of deep networks or graph structures, these hybrid approaches address the challenges of complex data distributions, offering scalable and robust solutions for diverse AD tasks. This synergy underscores the potential of such methods to push the boundaries of accuracy and adaptability in real-world applications.

### C. Support Vector Data Description

Support Vector Data Description (SVDD) [156] is a traditional machine learning method used to define a boundary around normal data points, effectively distinguishing them from anomalies. Unlike binary classification, SVDD is particularly effective for one-class classification tasks, where only normal data is available. This approach is computationally efficient and interpretable, as it provides a clear boundary between normal and abnormal points. By integrating SVDD with deep learning, researchers can enhance the boundary definition based on high-dimensional features extracted by a neural network, resulting in a model that combines the boundary precision of SVDD with the feature richness of deep learning. This hybrid model is highly effective in scenarios where boundary clarity and interpretability are paramount, such as in industrial monitoring or fraud detection.

To improve latent representations, Zhou *et al.* [169] propose Deep SVDD-VAE, which jointly optimizes VAE and SVDD. The VAE reconstructs input data, and SVDD simultaneously defines a spherical boundary in the latent space, ensuring separability of normal and anomalous instances. This joint optimization significantly outperforms traditional AE-based methods, as shown on MNIST, CIFAR-10, and GTSRB datasets. For variable-length time series data, Ergen *et al.* [124] introduce an LSTM-based AD framework, where LSTM and SVDD are jointly optimized using modified objectives. This method extends seamlessly to GRU architectures, demonstrating strong performance across unsupervised, semisupervised, and supervised settings. Besides, Zhang *et al.* [170] propose Deep Structure Preservation SVDD (DSPSVDD), which simultaneously minimizes hypersphere volume and network reconstruction error. This dual objective ensures deep feature preservation and enhances AD performance, outperforming traditional SVDD models on datasets like MNIST and MVTec AD.

These studies highlight the strengths of combining SVDD with deep learning, where deep models enhance feature representation while SVDD ensures boundary precision. This hybrid framework effectively addresses limitations in both methods, offering a scalable and interpretable solution for complex AD tasks across diverse domains.

### D. Summary and Insights

This section explores the integration of traditional and deep learning methods for AD, highlighting how their complementary strengths can be combined. Traditional methods, known for their simplicity, interpretability, and computational efficiency, excel in scenarios where transparency is critical. In contrast, deep learning methods offer superior adaptability

and accuracy, particularly for high-dimensional and unstructured data. By integrating these approaches, hybrid models can leverage the interpretability of traditional methods while retaining the robustness and flexibility of deep learning. This fusion not only enhances AD performance but also bridges the gap between accuracy and model transparency, making it a promising direction for future research.

## VI. OPEN ISSUES AND FUTURE WORKS

### A. Data Collection

Data scarcity and class imbalance remain major challenges in AD. Since anomalies are rare, obtaining large labeled datasets is costly and time-consuming, especially when expert annotation is required. Supervised learning struggles due to the lack of abnormal samples, while the overwhelming presence of normal data biases models toward common patterns. This problem is particularly critical in cybersecurity, healthcare, and industrial monitoring, where undetected anomalies can have serious consequences.

Several approaches mitigate these issues. Semi-supervised and unsupervised learning exploit normal data distributions to detect deviations without requiring labeled anomalies [171] [172]. Data augmentation, synthetic data generation, and oversampling improve data balance by increasing the number of anomalous examples, helping models generalize better [173] [174]. Despite these advancements, challenges remain. Semi-supervised methods struggle with subtle anomalies that closely resemble normal data. Augmentation techniques, often based on simple transformations, may fail to capture complex domain-specific variations. Similarly, synthetic data generation may not fully reflect real-world anomaly diversity, leading to models biased toward normal samples. Moreover, even with augmentation, models risk overfitting to the majority class, compromising anomaly detection performance. Ensuring that models remain sensitive to rare anomalies while maintaining accuracy on normal data remains an ongoing challenge. Future research may focus on refining self-supervised learning [175], improving the diversity of synthetic samples [176], and developing more adaptive anomaly detection frameworks to enhance robustness in real-world applications.

### B. Computational Complexity

In AD, computational complexity is a crucial factor, especially for systems operating in real-time environments or handling large-scale datasets. The efficiency of an algorithm directly impacts its feasibility in fields like industrial monitoring, cybersecurity, and autonomous systems, where swift detection is essential. Many advanced models, particularly deep learning approaches like autoencoders, GANs, and LSTMs, are computationally intensive due to their complex architectures and iterative learning processes. This often leads to trade-offs between detection accuracy and computational efficiency, with continuous efforts aimed at optimizing models to reduce computational demands without sacrificing performance.

Moreover, AD models frequently require substantial memory resources, especially when dealing with high-dimensional

or streaming data, making memory usage a crucial consideration. Techniques like memory-efficient architectures, data compression, and sparse modeling are commonly used to address this issue. Real-time AD adds further complexity, as algorithms must process incoming data and make rapid decisions in applications like autonomous driving and fraud detection [177], where even minimal delays can have severe consequences. Achieving real-time performance typically involves optimizing data processing speeds and decision-making through lightweight models [178] [179] and parallel processing techniques, such as GPU acceleration [180]. However, balancing real-time detection capabilities with high accuracy remains challenging.

The tension between computational complexity and detection accuracy persists, as complex models often excel in detection but lack practical applicability for real-time or large-scale scenarios. Simpler models, though computationally efficient, may fail to detect nuanced anomalies. Hybrid models or multi-stage frameworks that deploy complex methods only as needed provide a potential solution. Additionally, future research may benefit from exploring distributed computing solutions, like cloud [181] or edge computing, to enhance real-time AD performance in resource-limited environments.

### C. Explainability and Interpretability

Deep learning methods have greatly advanced AD by capturing complex patterns in high-dimensional data. However, they are often criticized as “black-box” models due to their lack of transparency, making it challenging to understand why certain data points are flagged as anomalies. For fields like healthcare, finance, or industrial monitoring, accurate detection alone is insufficient; stakeholders also need clear explanations to understand why a particular anomaly was detected. This lack of interpretability limits the practical deployment of deep learning models, as the inability to justify decisions reduces trust and hinders adoption in critical applications.

In fields like healthcare, where anomalies may be linked to medical diagnoses, or in finance, where fraud detection can carry legal implications, interpretability is essential. Transparent model decisions enable experts to validate results and make informed decisions. In safety-critical applications, such as autonomous driving or industrial equipment monitoring, understanding the rationale behind AD is vital for ensuring safety. One major challenge is balancing the trade-off between model interpretability and performance. Simpler models, like decision trees or linear regression, offer greater transparency but often lack the complexity needed to detect subtle anomalies in high-dimensional data. In contrast, deep learning models provide high accuracy but are harder to interpret.

Ongoing research is exploring hybrid approaches, where interpretable models are combined with more complex ones, allowing for accurate AD with the added benefit of interpretability. For example, attention mechanisms [182] in neural networks can help highlight specific data regions influencing decisions, providing insights into the model’s internal workings. Alternatively, tools like Local Interpretable Model-agnostic Explanations (LIME) and SHapley Additive exPlanations (SHAP) [2] can offer post-hoc explanations, improving

transparency without altering model structure. Future research could also focus on real-time explainability in time-sensitive applications, and incorporating domain knowledge or user feedback to enhance model interpretability.

#### D. Handling Diverse Types of Anomalies

In real-world AD, multiple types of anomalies often coexist, adding complexity to the detection process. Beyond point anomalies, which are the simplest, other types like contextual and collective anomalies are common, especially in dynamic environments. For instance, in intelligent transportation systems, anomalies may include both isolated incidents (e.g., a single vehicle's sudden deceleration) and collective patterns (e.g., multiple vehicles simultaneously slowing down), each requiring different detection methods. Effectively capturing these varied anomaly types requires flexible models capable of adapting to different anomaly patterns without focusing on only one type.

Continuous research is needed to develop models that can generalize across anomaly types, enhancing adaptability and balancing detection accuracy with model flexibility. Hybrid approaches, for instance, can integrate different methods to capture diverse anomalies more effectively. The challenge remains in achieving this versatility without sacrificing accuracy, as models must maintain strong performance across different contexts. Future work may also explore multi-modal models [183] that combine different types of data, further improving detection capabilities by drawing from diverse data sources. These directions aim to create AD systems that are both robust and adaptable, capable of handling the complex and mixed nature of real-world anomaly scenarios.

## VII. CONCLUSION

In this survey, we have provided a comprehensive overview of the recent advancements in AD with a primary focus on deep learning techniques from 2019 to 2024. By analyzing over 180 research papers from leading journals and conferences, we have explored how AD methods have evolved to address diverse challenges across various types of data. This survey categorizes and examines deep learning methods into reconstruction-based, prediction-based, and hybrid approaches, highlighting their strengths, limitations, and applications. Recognizing the simplicity, interpretability, and computational efficiency of traditional AD methods, we reviewed their integration with deep learning techniques. These hybrid approaches aim to leverage the strengths of both paradigms, enhancing robustness and efficiency in AD systems. This survey not only sheds light on the state-of-the-art techniques but also identifies gaps and opportunities for future research. By focusing on the latest trends and innovations, this work aims to inspire further exploration and advancements in the rapidly evolving field of AD.

## REFERENCES

- [1] L. Ruff, J. R. Kauffmann, R. A. Vandermeulen, G. Montavon, W. Samek, M. Kloft, T. G. Dietterich, and K.-R. Müller, "A unifying review of deep and shallow anomaly detection," *Proceed. IEEE*, vol. 109, no. 5, pp. 756–795, 2021.
- [2] V. Vimbi, N. Shaffi, and M. Mahmud, "Interpreting artificial intelligence models: a systematic review on the application of lime and shap in alzheimer's disease detection," *Brain Informatics*, vol. 11, no. 1, p. 10, 2024.
- [3] F. Al-Turjman, H. Zahmatkesh, and R. Shahroze, "An overview of security and privacy in smart cities' iot communications," *Trans. Emerg. Telecommun. Technol.*, vol. 33, no. 3, p. e3677, 2022.
- [4] Y. A. Qadri, A. Nauman, Y. B. Zikria, A. V. Vasilakos, and S. W. Kim, "The future of healthcare internet of things: a survey of emerging technologies," *IEEE Commun. Surv. Tutor.*, vol. 22, no. 2, pp. 1121–1167, 2020.
- [5] M. Humayun, N. Jhanjhi, B. Hamid, and G. Ahmed, "Emerging smart logistics and transportation using iot and blockchain," *IEEE Internet Things Mag.*, vol. 3, no. 2, pp. 58–62, 2020.
- [6] S. H. Haji and S. Y. Ameen, "Attack and anomaly detection in iot networks using machine learning techniques: A review," *Asian J. Res. Comput. Sci.*, vol. 9, no. 2, pp. 30–46, 2021.
- [7] V. Mothukuri, P. Khare, R. M. Parizi, S. Pouriyeh, A. Dehghantanha, and G. Srivastava, "Federated-learning-based anomaly detection for iot security attacks," *IEEE Internet Things J.*, vol. 9, no. 4, pp. 2545–2554, 2021.
- [8] S. A. Al Mamun and J. Valimaki, "Anomaly detection and classification in cellular networks using automatic labeling technique for applying supervised learning," *Procedia Comput. Sci.*, vol. 140, pp. 186–195, 2018.
- [9] M. E. Villa-Pérez, M. A. Alvarez-Carmona, O. Loyola-Gonzalez, M. A. Medina-Pérez, J. C. Velasco-Rossell, and K.-K. R. Choo, "Semi-supervised anomaly detection algorithms: A comparative summary and future research directions," *Knowledge-Based Systems*, vol. 218, p. 106878, 2021.
- [10] G. Michau and O. Fink, "Unsupervised transfer learning for anomaly detection: Application to complementary operating condition transfer," *Knowledge-Based Systems*, vol. 216, p. 106816, 2021.
- [11] Y. Liang, J. Zhang, S. Zhao, R. Wu, Y. Liu, and S. Pan, "Omnifrequency channel-selection representations for unsupervised anomaly detection," *IEEE Trans. Image Process.*, 2023.
- [12] B. Siegel, "Industrial anomaly detection: A comparison of unsupervised neural network architectures," *IEEE Sens. Lett.*, vol. 4, no. 8, pp. 1–4, 2020.
- [13] P. Bergmann, M. Fauser, D. Sattlegger, and C. Steger, "Mvtac ad—a comprehensive real-world dataset for unsupervised anomaly detection," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 9592–9600.
- [14] S. Schmidl, P. Wenig, and T. Papenbrock, "Anomaly detection in time series: a comprehensive evaluation," *Proc. VLDB Endow.*, vol. 15, no. 9, pp. 1779–1797, 2022.
- [15] S. Zhai, Y. Cheng, W. Lu, and Z. Zhang, "Deep structured energy based models for anomaly detection," in *Int. Conf. Mach. Learn. (ICML)*. PMLR, 2016, pp. 1100–1109.
- [16] H. Sarmadi and A. Karamodin, "A novel anomaly detection method based on adaptive mahalanobis-squared distance and one-class knn rule for structural health monitoring under environmental effects," *Mech. Syst. Signal Process.*, vol. 140, p. 106495, 2020.
- [17] I. Syarif, A. Prugel-Bennett, and G. Wills, "Unsupervised clustering approach for network anomaly detection," in *Netw. Digit. Technol., Int. Conf., NDT 2012, Proc., Part I*. Springer, 2012, pp. 135–145.
- [18] D. Samariya and A. Thakkar, "A comprehensive survey of anomaly detection algorithms," *Ann. Data Sci.*, vol. 10, no. 3, pp. 829–850, 2023.
- [19] G. Pang, C. Shen, L. Cao, and A. V. D. Hengel, "Deep learning for anomaly detection: A review," *ACM Comput. Surv.*, vol. 54, no. 2, pp. 1–38, 2021.
- [20] L. Bergman, N. Cohen, and Y. Hoshen, "Deep nearest neighbor anomaly detection," *arXiv preprint arXiv:2002.10445*, 2020.
- [21] K. Leung and C. Leckie, "Unsupervised anomaly detection in network intrusion detection using clusters," in *Proc. 28th Australas. Conf. Comput. Sci.*, vol. 38, 2005, pp. 333–342.
- [22] H. Ringberg, A. Soule, J. Rexford, and C. Diot, "Sensitivity of pca for traffic anomaly detection," in *Proc. 2007 ACM SIGMETRICS Int. Conf. Meas. Model. Comput. Syst.*, 2007, pp. 109–120.
- [23] D. Kwon, H. Kim, J. Kim, S. C. Suh, I. Kim, and K. J. Kim, "A survey of deep learning-based network anomaly detection," *Cluster Computing*, vol. 22, pp. 949–961, 2019.
- [24] A. Aldweesh, A. Derhab, and A. Z. Emam, "Deep learning approaches for anomaly-based intrusion detection systems: A survey, taxonomy, and open issues," *Knowl.-Based Syst.*, vol. 189, p. 105124, 2020.

- [25] L. Li, J. Yan, H. Wang, and Y. Jin, "Anomaly detection of time series with smoothness-inducing sequential variational auto-encoder," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 3, pp. 1177–1191, 2020.
- [26] G. Harshvardhan, M. K. Gourisaria, M. Pandey, and S. S. Rautaray, "A comprehensive survey and analysis of generative models in machine learning," *Comput. Sci. Rev.*, vol. 38, p. 100285, 2020.
- [27] B. Nachman and D. Shih, "Anomaly detection with density estimation," *Phys. Rev. D*, vol. 101, no. 7, p. 075042, 2020.
- [28] A. B. Nassif, M. A. Talib, Q. Nasir, and F. M. Dakalbab, "Machine learning for anomaly detection: A systematic review," *IEEE Access*, vol. 9, pp. 78 658–78 700, 2021.
- [29] X. Ma, J. Wu, S. Xue, J. Yang, C. Zhou, Q. Z. Sheng, H. Xiong, and L. Akoglu, "A comprehensive survey on graph anomaly detection with deep learning," *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 12, pp. 12 012–12 038, 2021.
- [30] X. Xia, X. Pan, N. Li, X. He, L. Ma, X. Zhang, and N. Ding, "Gan-based anomaly detection: A review," *Neurocomputing*, vol. 493, pp. 497–535, 2022.
- [31] J. Lv, Y. Wang, and S. Chen, "Adaptive multivariate time-series anomaly detection," *Inf. Process. Manag.*, vol. 60, no. 4, p. 103383, 2023.
- [32] M. Y. I. Basheer, A. M. Ali, N. H. A. Hamid, M. A. M. Ariffin, R. Osman, S. Nordin, and X. Gu, "Autonomous anomaly detection for streaming data," *Knowledge-Based Systems*, vol. 284, p. 111235, 2024.
- [33] X. Peng, H. Li, F. Yuan, S. G. Razul, Z. Chen, and Z. Lin, "An extreme learning machine for unsupervised online anomaly detection in multivariate time series," *Neurocomputing*, vol. 501, pp. 596–608, 2022.
- [34] Y. Choi, H. Lim, H. Choi, and I.-J. Kim, "Gan-based anomaly detection and localization of multivariate time series data for power plant," in *Proc. 2020 IEEE Int. Conf. Big Data Smart Comput. (BigComp)*. IEEE, 2020, pp. 71–74.
- [35] H.-T. Duong, V.-T. Le, and V. T. Hoang, "Deep learning-based anomaly detection in video surveillance: a survey," *Sensors*, vol. 23, no. 11, p. 5024, 2023.
- [36] S. Thudumu, P. Branch, J. Jin, and J. Singh, "A comprehensive survey of anomaly detection techniques for high dimensional big data," *Journal of Big Data*, vol. 7, pp. 1–30, 2020.
- [37] I. Souiden, M. N. Omri, and Z. Brahmi, "A survey of outlier detection in high dimensional data streams," *Comput. Sci. Rev.*, vol. 44, p. 100463, 2022.
- [38] Q. Ding and E. D. Kolaczyk, "A compressed pca subspace method for anomaly detection in high-dimensional data," *IEEE Trans. Inf. Theory*, vol. 59, no. 11, pp. 7419–7433, 2013.
- [39] M. Sakurada and T. Yairi, "Anomaly detection using autoencoders with nonlinear dimensionality reduction," in *Proc. MLSDA 2014 2nd Workshop Mach. Learn. Sensory Data Anal.*, 2014, pp. 4–11.
- [40] T. Cheng and B. Wang, "Total variation and sparsity regularized decomposition model with union dictionary for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 2, pp. 1472–1486, 2020.
- [41] L. Li, W. Li, Q. Du, and R. Tao, "Low-rank and sparse decomposition with mixture of gaussian for hyperspectral anomaly detection," *IEEE Trans. Cybern.*, vol. 51, no. 9, pp. 4363–4372, 2021.
- [42] S. Han and S. S. Woo, "Learning sparse latent graph representations for anomaly detection in multivariate time series," in *Proc. 28th ACM SIGKDD Conf. Knowl. Discov. Data Min.*, 2022, pp. 2977–2986.
- [43] X. Ma and W. Shi, "Aesmote: Adversarial reinforcement learning with smote for anomaly detection," *IEEE Trans. Netw. Sci. Eng.*, vol. 8, no. 2, pp. 943–956, 2021.
- [44] M. Kim, E. Ou, P.-L. Loh, T. Allen, R. Agasie, and K. Liu, "Rnn-based online anomaly detection in nuclear reactors for highly imbalanced datasets with uncertainty," *Nucl. Eng. Des.*, vol. 364, p. 110699, 2020.
- [45] G. Dlamini and M. Fahim, "Dgm: a data generative model to improve minority class presence in anomaly detection domain," *Neural Comput. Appl.*, vol. 33, pp. 13 635–13 646, 2021.
- [46] S. Han, X. Hu, H. Huang, M. Jiang, and Y. Zhao, "Adbench: Anomaly detection benchmark," *Adv. Neural Inf. Process. Syst.*, vol. 35, pp. 32 142–32 159, 2022.
- [47] Y. Zhang, Y. Chen, J. Wang, and Z. Pan, "Unsupervised deep anomaly detection for multi-sensor time-series signals," *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 2, pp. 2118–2132, 2023.
- [48] D. Chen, L. Yue, X. Chang, M. Xu, and T. Jia, "Nm-gan: Noise-modulated generative adversarial network for video anomaly detection," *Pattern Recognition*, vol. 116, p. 107969, 2021.
- [49] M. U. Hassan, M. H. Rehmani, and J. Chen, "Anomaly detection in blockchain networks: A comprehensive survey," *IEEE Commun. Surv. Tutor.*, vol. 25, no. 1, pp. 289–318, 2022.
- [50] Y. Liu, S. Garg, J. Nie, Y. Zhang, Z. Xiong, J. Kang, and M. S. Hossain, "Deep anomaly detection for time-series data in industrial iot: A communication-efficient on-device federated learning approach," *IEEE Internet Things J.*, vol. 8, no. 8, pp. 6348–6358, 2020.
- [51] M. J. Idrissi, H. Alami, A. El Mahdaoui, A. El Mekki, S. Oualil, Z. Yartaoui, and I. Berrada, "Fed-anids: Federated learning for anomaly-based network intrusion detection systems," *Expert Syst. Appl.*, vol. 234, p. 121000, 2023.
- [52] L. Cui, Y. Qu, G. Xie, D. Zeng, R. Li, S. Shen, and S. Yu, "Security and privacy-enhanced federated learning for anomaly detection in iot infrastructures," *IEEE Trans. Ind. Inform.*, vol. 18, no. 5, pp. 3492–3500, 2022.
- [53] X. Wang, J. Liu, T. Qiu, C. Mu, C. Chen, and P. Zhou, "A real-time collision prediction mechanism with deep learning for intelligent transportation system," *IEEE Trans. Veh. Technol.*, vol. 69, no. 9, pp. 9497–9508, 2020.
- [54] G. Li, T.-H. Nguyen, and J. J. Jung, "Traffic incident detection based on dynamic graph embedding in vehicular edge computing," *Appl. Sci.*, vol. 11, no. 13, p. 5861, 2021.
- [55] G. Li and J. J. Jung, "Deep learning for anomaly detection in multivariate time series: Approaches, applications, and challenges," *Inf. Fusion*, vol. 91, pp. 93–102, 2023.
- [56] C. Zhao, X. Chang, T. Xie, H. Fujita, and J. Wu, "Unsupervised anomaly detection based method of risk evaluation for road traffic accident," *Appl. Intell.*, vol. 53, no. 1, pp. 369–384, 2023.
- [57] S. Li, A. Pandey, B. Hooi, C. Faloutsos, and L. Pileggi, "Dynamic graph-based anomaly detection in the electrical grid," *IEEE Trans. Power Syst.*, vol. 37, no. 5, pp. 3408–3422, 2022.
- [58] X. Wang and S.-H. Ahn, "Real-time prediction and anomaly detection of electrical load in a residential community," *Appl. Energy*, vol. 259, p. 114145, 2020.
- [59] I. Sinosoglou, P. Radoglou-Grammatikis, G. Efstathopoulos, P. Foullas, and P. Sarigiannidis, "A unified deep learning anomaly detection and classification approach for smart grid environments," *IEEE Trans. Netw. Serv. Manag.*, vol. 18, no. 2, pp. 1137–1151, 2021.
- [60] T. Fernando, H. Gammulle, S. Denman, S. Sridharan, and C. Fookes, "Deep learning for medical anomaly detection—a survey," *ACM Comput. Surv.*, vol. 54, no. 7, pp. 1–37, 2021.
- [61] E. Šabić, D. Keeley, B. Henderson, and S. Nannemann, "Healthcare and anomaly detection: using machine learning to predict anomalies in heart rate data," *Ai & Society*, vol. 36, no. 1, pp. 149–158, 2021.
- [62] K. G. Al-Hashedi and P. Magalingam, "Financial fraud detection applying data mining techniques: A comprehensive review from 2009 to 2019," *Comput. Sci. Rev.*, vol. 40, p. 100402, 2021.
- [63] W. Hilal, S. A. Gadsden, and J. Yawney, "Financial fraud: a review of anomaly detection techniques and recent advances," *Expert Syst. Appl.*, vol. 193, p. 116429, 2022.
- [64] H. Fujita, A. Gaeta, V. Loia, and F. Orciuoli, "Resilience analysis of critical infrastructures: A cognitive approach based on granular computing," *IEEE Trans. Cybern.*, vol. 49, no. 5, pp. 1835–1848, 2019.
- [65] V. K. Singh and M. Govindarasu, "A cyber-physical anomaly detection for wide-area protection using machine learning," *IEEE Trans. Smart Grid*, vol. 12, no. 4, pp. 3514–3526, 2021.
- [66] S. M. Nagarajan, G. G. Devarajan, A. K. Bashir, R. P. Mahapatra, and M. S. Al-Numay, "Iadf-cps: Intelligent anomaly detection framework towards cyber physical systems," *Comput. Commun.*, vol. 188, pp. 81–89, 2022.
- [67] T. Nakao, S. Hanaoka, Y. Nomura, M. Murata, T. Takenaga, S. Miki, T. Watadani, T. Yoshikawa, N. Hayashi, and O. Abe, "Unsupervised deep anomaly detection in chest radiographs," *J. Digit. Imaging*, vol. 34, pp. 418–427, 2021.
- [68] W. H. Pinaya, P.-D. Tudosiu, R. Gray, G. Rees, P. Nachev, S. Ourselin, and M. J. Cardoso, "Unsupervised brain imaging 3d anomaly detection and segmentation with transformers," *Med. Image Anal.*, vol. 79, p. 102475, 2022.
- [69] L. Chen, Z. You, N. Zhang, J. Xi, and X. Le, "Utrad: Anomaly detection and localization with u-transformer," *Neural Networks*, vol. 147, pp. 53–62, 2022.
- [70] N. Shvetsova, B. Bakker, I. Fedulova, H. Schulz, and D. V. Dylov, "Anomaly detection in medical imaging with deep perceptual autoencoders," *IEEE Access*, vol. 9, pp. 118 571–118 583, 2021.
- [71] R. L. Draelos, D. Dov, M. A. Mazurowski, J. Y. Lo, R. Henao, G. D. Rubin, and L. Carin, "Machine-learning-based multiple abnormality

- prediction with large-scale chest computed tomography volumes," *Med. Image Anal.*, vol. 67, p. 101857, 2021.
- [72] N. Shvetsova, B. Bakker, I. Fedulova, H. Schulz, and D. V. Dylov, "Anomaly detection in medical imaging with deep perceptual autoencoders," *IEEE Access*, vol. 9, pp. 118 571–118 583, 2021.
- [73] H. Zhao, Y. Li, N. He, K. Ma, L. Fang, H. Li, and Y. Zheng, "Anomaly detection for medical images using self-supervised and translation-consistent features," *IEEE Trans. Med. Imaging*, vol. 40, no. 12, pp. 3641–3651, 2021.
- [74] R. Nayak, U. C. Pati, and S. K. Das, "A comprehensive review on deep learning-based methods for video anomaly detection," *Image Vis. Comput.*, vol. 106, p. 104078, 2021.
- [75] Y. Wang, T. Liu, J. Zhou, and J. Guan, "Video anomaly detection based on spatio-temporal relationships among objects," *Neurocomputing*, vol. 532, pp. 141–151, 2023.
- [76] N. Li, F. Chang, and C. Liu, "Spatial-temporal cascade autoencoder for video anomaly detection in crowded scenes," *IEEE Trans. Multimed.*, vol. 23, pp. 203–215, 2020.
- [77] D. Chen, P. Wang, L. Yue, Y. Zhang, and T. Jia, "Anomaly detection in surveillance video based on bidirectional prediction," *Image Vis. Comput.*, vol. 98, p. 103915, 2020.
- [78] M. H. Bhuyan, D. K. Bhattacharyya, and J. K. Kalita, "Network anomaly detection: methods, systems and tools," *IEEE Commun. Surv. Tutor.*, vol. 16, no. 1, pp. 303–336, 2013.
- [79] S. Liu, B. Zhou, Q. Ding, B. Hooi, Z. Zhang, H. Shen, and X. Cheng, "Time series anomaly detection with adversarial reconstruction networks," *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 4, pp. 4293–4306, 2022.
- [80] H. Cao, C. Tan, Z. Gao, Y. Xu, G. Chen, P.-A. Heng, and S. Z. Li, "A survey on generative diffusion models," *IEEE Transactions on Knowledge and Data Engineering*, 2024.
- [81] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," *Commun. ACM*, vol. 63, no. 11, pp. 139–144, 2020.
- [82] L. Yang, Z. Zhang, Y. Song, S. Hong, R. Xu, Y. Zhao, W. Zhang, B. Cui, and M.-H. Yang, "Diffusion models: A comprehensive survey of methods and applications," *ACM Comput. Surv.*, vol. 56, no. 4, pp. 1–39, 2023.
- [83] S. Bond-Taylor, A. Leach, Y. Long, and C. G. Willcocks, "Deep generative modelling: A comparative review of vae's, gans, normalizing flows, energy-based and autoregressive models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 11, pp. 7327–7347, 2021.
- [84] S. Sheynin, S. Benaim, and L. Wolf, "A hierarchical transformation-discriminating generative model for few shot anomaly detection," in *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, 2021, pp. 8495–8504.
- [85] W. Lim, K. Y. S. Chek, L. B. Theng, and C. T. C. Lin, "Future of generative adversarial networks (gan) for anomaly detection in network security: A review," *Comput. Secur.*, p. 103733, 2024.
- [86] X. Du, J. Chen, J. Yu, S. Li, and Q. Tan, "Generative adversarial nets for unsupervised outlier detection," *Expert Syst. Appl.*, vol. 236, p. 121161, 2024.
- [87] J. Wu, Z. Zhao, C. Sun, R. Yan, and X. Chen, "Fault-attention generative probabilistic adversarial autoencoder for machine anomaly detection," *IEEE Trans. Ind. Inf.*, vol. 16, no. 12, pp. 7479–7488, 2020.
- [88] F. Dong, Y. Zhang, and X. Nie, "Dual discriminator generative adversarial network for video anomaly detection," *IEEE Access*, vol. 8, pp. 88 170–88 176, 2020.
- [89] D. Guo, Z. Liu, and R. Li, "Regraphgan: A graph generative adversarial network model for dynamic network anomaly detection," *Neural Networks*, vol. 166, pp. 273–285, 2023.
- [90] Y. Liu, Z. Li, C. Zhou, Y. Jiang, J. Sun, M. Wang, and X. He, "Generative adversarial active learning for unsupervised outlier detection," *IEEE Trans. Knowl. Data Eng.*, vol. 32, no. 8, pp. 1517–1528, 2019.
- [91] C. Liu, Z. Kong, S. Babu, C. Joslin, and J. Ferguson, "An integrated manifold learning approach for high-dimensional data feature extractions and its applications to online process monitoring of additive manufacturing," *IIE Transactions*, vol. 53, no. 11, pp. 1215–1230, 2021.
- [92] J. Miao, H. Tao, H. Xie, J. Sun, and J. Cao, "Reconstruction-based anomaly detection for multivariate time series using contrastive generative adversarial networks," *Inf. Process. Manag.*, vol. 61, no. 1, p. 103569, 2024.
- [93] R. Liu, W. Liu, Z. Zheng, L. Wang, L. Mao, Q. Qiu, and G. Ling, "Anomaly-gan: A data augmentation method for train surface anomaly detection," *Expert Syst. Appl.*, vol. 228, p. 120284, 2023.
- [94] Y. Li, Z. Shi, C. Liu, W. Tian, Z. Kong, and C. B. Williams, "Augmented time regularized generative adversarial network (atr-gan) for data augmentation in online process anomaly detection," *IEEE Trans. Autom. Sci. Eng.*, vol. 19, no. 4, pp. 3338–3355, 2021.
- [95] L. Zhang, W. Bai, X. Xie, L. Chen, and P. Dong, "Tmanomaly: Time-series mutual adversarial networks for industrial anomaly detection," *IEEE Trans. Ind. Inform.*, 2023.
- [96] B. Du, X. Sun, J. Ye, K. Cheng, J. Wang, and L. Sun, "Gan-based anomaly detection for multivariate time series using polluted training set," *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 12, pp. 12 208–12 219, 2021.
- [97] G. Fan, Y. Ma, X. Mei, F. Fan, J. Huang, and J. Ma, "Hyperspectral anomaly detection with robust graph autoencoders," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–14, 2021.
- [98] S. Wang, X. Wang, L. Zhang, and Y. Zhong, "Auto-ad: Autonomous hyperspectral anomaly detection network based on fully convolutional autoencoder," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–14, 2021.
- [99] H. Liu, X. Su, X. Shen, and X. Zhou, "Msnet: Self-supervised multi-scale network with enhanced separation training for hyperspectral anomaly detection," *IEEE Trans. Geosci. Remote Sens.*, 2024.
- [100] X. Lin, Z. Li, H. Fan, Y. Fu, and X. Chen, "Exploiting negative correlation for unsupervised anomaly detection in contaminated time series," *Expert Syst. Appl.*, p. 123535, 2024.
- [101] C. Huang, Z. Yang, J. Wen, Y. Xu, Q. Jiang, J. Yang, and Y. Wang, "Self-supervision-augmented deep autoencoder for unsupervised visual anomaly detection," *IEEE Trans. Cybern.*, vol. 52, no. 12, pp. 13 834–13 847, 2021.
- [102] C. Yin, S. Zhang, J. Wang, and N. N. Xiong, "Anomaly detection based on convolutional recurrent autoencoder for iot time series," *IEEE Trans. Syst. Man Cybern.: Syst.*, vol. 52, no. 1, pp. 112–122, 2020.
- [103] W. Zhang, C. Zhang, and F. Tsung, "Grelen: Multivariate time series anomaly detection from the perspective of graph relational learning," in *IJCAI*, 2022, pp. 2390–2397.
- [104] X. Zhou, Y. Hu, W. Liang, J. Ma, and Q. Jin, "Variational lstm enhanced anomaly detection for industrial big data," *IEEE Trans. Ind. Inform.*, vol. 17, no. 5, pp. 3469–3477, 2020.
- [105] A. Makhzani, J. Shlens, N. Jaitly, I. Goodfellow, and B. Frey, "Adversarial autoencoders," *arXiv preprint arXiv:1511.05644*, 2015.
- [106] Q. Su, B. Tian, H. Wan, and J. Yin, "Anomaly detection under contaminated data with contamination-immune bidirectional gans," *IEEE Trans. Knowl. Data Eng.*, 2024.
- [107] J. Yu, X. Gao, F. Zhai, B. Li, B. Xue, S. Fu, L. Chen, and Z. Meng, "An adversarial contrastive autoencoder for robust multivariate time series anomaly detection," *Expert Syst. Appl.*, vol. 245, p. 123010, 2024.
- [108] J. Ho, A. Jain, and P. Abbeel, "Denoising diffusion probabilistic models," *Adv. Neural Inf. Process. Syst.*, vol. 33, pp. 6840–6851, 2020.
- [109] J. Wolleb, F. Bieder, R. Sandkuhler, and P. C. Cattin, "Diffusion models for medical anomaly detection," in *Int. Conf. Med. Image Comput. Comput.-Assist. Interv. (MICCAI)*. Springer, 2022, pp. 35–45.
- [110] X. Zhang, N. Li, J. Li, T. Dai, Y. Jiang, and S.-T. Xia, "Unsupervised surface anomaly detection with diffusion probabilistic model," in *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, 2023, pp. 6782–6791.
- [111] S. Li, J. Yu, Y. Lu, G. Yang, X. Du, and S. Liu, "Self-supervised enhanced denoising diffusion for anomaly detection," *Inf. Sci.*, vol. 669, p. 120612, 2024.
- [112] J. Zeng, X. Liu, and Z. Li, "Radio anomaly detection based on improved denoising diffusion probabilistic models," *IEEE Commun. Lett.*, 2023.
- [113] X. Li, C. Xiao, Z. Feng, S. Pang, W. Tai, and F. Zhou, "Controlled graph neural networks with denoising diffusion for anomaly detection," *Expert Syst. Appl.*, vol. 237, p. 121533, 2024.
- [114] C. Li, G. Feng, Y. Li, R. Liu, Q. Miao, and L. Chang, "Diffatad: Denoising diffusion probabilistic models for vehicle trajectory anomaly detection," *Knowledge-Based Systems*, vol. 286, p. 111387, 2024.
- [115] J. Pei, J. Wang, D. Shi, and P. Wang, "Detection and imputation-based two-stage denoising diffusion power system measurement recovery under cyber-physical uncertainties," *IEEE Trans. Smart Grid*, vol. 15, no. 6, pp. 5965–5980, 2024.
- [116] H. He, J. Zhang, H. Chen, X. Chen, Z. Li, X. Chen, Y. Wang, C. Wang, and L. Xie, "A diffusion-based framework for multi-class anomaly detection," in *Proc. AAAI Conf. Artif. Intell.*, vol. 38, no. 8, 2024, pp. 8472–8480.
- [117] A. Sherstinsky, "Fundamentals of recurrent neural network (rnn) and long short-term memory (lstm) network," *Physica D: Nonlinear Phenomena*, vol. 404, p. 132306, 2020.

- [118] G. Van Houdt, C. Mosquera, and G. Nápoles, "A review on the long short-term memory model," *Artif. Intell. Rev.*, vol. 53, no. 8, pp. 5929–5955, 2020.
- [119] B. Lindemann, B. Maschler, N. Sahlab, and M. Weyrich, "A survey on anomaly detection for technical systems using lstm networks," *Comput. Ind.*, vol. 131, p. 103498, 2021.
- [120] R. Dey and F. M. Salem, "Gate-variants of gated recurrent unit (gru) neural networks," in *Proc. 2017 IEEE 60th Int. Midwest Symp. Circuits Syst. (MWSCAS)*. IEEE, 2017, pp. 1597–1600.
- [121] Y. Wang, X. Du, Z. Lu, Q. Duan, and J. Wu, "Improved lstm-based time-series anomaly detection in rail transit operation environments," *IEEE Trans. Ind. Inform.*, vol. 18, no. 12, pp. 9027–9036, 2022.
- [122] H. Chen, H. Liu, X. Chu, Q. Liu, and D. Xue, "Anomaly detection and critical scada parameters identification for wind turbines based on lstm-ae neural network," *Renew. Energy*, vol. 172, pp. 829–840, 2021.
- [123] P. Liu, X. Sun, Y. Han, Z. He, W. Zhang, and C. Wu, "Arrhythmia classification of lstm autoencoder based on time series anomaly detection," *Biomed. Signal Process. Control*, vol. 71, p. 103228, 2022.
- [124] Y. Yao, J. Ma, S. Feng, and Y. Ye, "Svd-ae: An asymmetric autoencoder with svd regularization for multivariate time series anomaly detection," *Neural Networks*, vol. 170, pp. 535–547, 2024.
- [125] S. Longari, D. H. N. Valcarcel, M. Zago, M. Carminati, and S. Zanero, "Cannano: An anomaly detection system based on lstm autoencoders for controller area network," *IEEE Trans. Netw. Serv. Manag.*, vol. 18, no. 2, pp. 1913–1924, 2020.
- [126] J. Pei, J. Wang, and D. Shi, "Data-driven measurement tampering detection considering spatial-temporal correlations," in *Proc. 2019 IEEE 3rd Conf. Energy Internet Energy Syst. Integr. (EI2)*, 2019, pp. 2641–2646.
- [127] T. Lei, C. Gong, G. Chen, M. Ou, K. Yang, and J. Li, "A novel unsupervised framework for time series data anomaly detection via spectrum decomposition," *Knowledge-Based Systems*, vol. 280, p. 111002, 2023.
- [128] D. Hu, S. Wu, J. Wang, and D. Shi, "Training a dynamic neural network to detect false data injection attacks under multiple unforeseen operating conditions," *IEEE Trans. Smart Grid*, 2023.
- [129] C. Tang, L. Xu, B. Yang, Y. Tang, and D. Zhao, "Gru-based interpretable multivariate time series anomaly detection in industrial control system," *Comput. Secur.*, vol. 127, p. 103094, 2023.
- [130] J. Yu, X. Gao, B. Li, F. Zhai, J. Lu, B. Xue, S. Fu, and C. Xiao, "A filter-augmented auto-encoder with learnable normalization for robust multivariate time series anomaly detection," *Neural Networks*, vol. 170, pp. 478–493, 2024.
- [131] A. Vaswani, "Attention is all you need," *Adv. Neural Inf. Process. Syst.*, 2017.
- [132] H. Kang and P. Kang, "Transformer-based multivariate time series anomaly detection using inter-variable attention mechanism," *Knowledge-Based Systems*, p. 111507, 2024.
- [133] J. Kim, H. Kang, and P. Kang, "Time-series anomaly detection with stacked transformer representations and 1d convolutional network," *Eng. Appl. Artif. Intell.*, vol. 120, p. 105964, 2023.
- [134] S. Tuli, G. Casale, and N. R. Jennings, "Tranad: Deep transformer networks for anomaly detection in multivariate time series data," *arXiv preprint arXiv:2201.07284*, 2022.
- [135] C. Wang and G. Liu, "From anomaly detection to classification with graph attention and transformer for multivariate time series," *Adv. Eng. Inform.*, vol. 60, p. 102357, 2024.
- [136] J. Fan, Z. Wang, H. Wu, D. Sun, J. Wu, and X. Lu, "An adversarial time-frequency reconstruction network for unsupervised anomaly detection," *Neural Networks*, vol. 168, pp. 44–56, 2023.
- [137] Y. Shi, B. Wang, Y. Yu, X. Tang, C. Huang, and J. Dong, "Robust anomaly detection for multivariate time series through temporal gcns and attention-based vae," *Knowledge-Based Systems*, vol. 275, p. 110725, 2023.
- [138] C. Ding, S. Sun, and J. Zhao, "Mst-gat: A multimodal spatial-temporal graph attention network for time series anomaly detection," *Inf. Fusion*, vol. 89, pp. 527–536, 2023.
- [139] W. Zhu, W. Li, E. R. Dorsey, and J. Luo, "Unsupervised anomaly detection by densely contrastive learning for time series data," *Neural Networks*, vol. 168, pp. 450–458, 2023.
- [140] H. Sun, M. Chen, J. Weng, Z. Liu, and G. Geng, "Anomaly detection for in-vehicle network using cnn-lstm with attention mechanism," *IEEE Trans. Veh. Technol.*, vol. 70, no. 10, pp. 10 880–10 893, 2021.
- [141] T. Le, H. C. Vu, A. Ponchet-Durupt, N. Boudaoud, Z. Cherfi-Boulanger, and T. Nguyen-Trang, "Unsupervised detecting anomalies in multivariate time series by robust convolutional lstm encoder-decoder (rcled)," *Neurocomputing*, vol. 592, p. 127791, 2024.
- [142] M. Jin, H. Y. Koh, Q. Wen, D. Zambon, C. Alippi, G. I. Webb, I. King, and S. Pan, "A survey on graph neural networks for time series: Forecasting, classification, imputation, and anomaly detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, 2024.
- [143] Y. Wu, H.-N. Dai, and H. Tang, "Graph neural networks for anomaly detection in industrial internet of things," *IEEE Internet Things J.*, vol. 9, no. 12, pp. 9214–9231, 2022.
- [144] H. Kim, B. S. Lee, W.-Y. Shin, and S. Lim, "Graph anomaly detection with graph neural networks: Current status and challenges," *IEEE Access*, vol. 10, pp. 111 820–111 829, 2022.
- [145] A. Deng and B. Hooi, "Graph neural network-based anomaly detection in multivariate time series," in *Proc. AAAI Conf. Artif. Intell. (AAAI)*, vol. 35, no. 5, 2021, pp. 4027–4035.
- [146] Z. Chen, D. Chen, X. Zhang, Z. Yuan, and X. Cheng, "Learning graph structures with transformer for multivariate time-series anomaly detection in iot," *IEEE Internet Things J.*, vol. 9, no. 12, pp. 9179–9189, 2021.
- [147] Y. Zheng, H. Y. Koh, M. Jin, L. Chi, K. T. Phan, S. Pan, Y.-P. P. Chen, and W. Xiang, "Correlation-aware spatial-temporal graph learning for multivariate time-series anomaly detection," *IEEE Trans. Neural Netw. Learn. Syst.*, 2023.
- [148] Y. Liu, Z. Li, S. Pan, C. Gong, C. Zhou, and G. Karypis, "Anomaly detection on attributed networks via contrastive self-supervised learning," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 6, pp. 2378–2392, 2022.
- [149] H. Zhao, Y. Wang, J. Duan, C. Huang, D. Cao, Y. Tong, B. Xu, J. Bai, J. Tong, and Q. Zhang, "Multivariate time-series anomaly detection via graph attention network," in *Proc. 2020 IEEE Int. Conf. Data Min. (ICDM)*. IEEE, 2020, pp. 841–850.
- [150] W. Chen, L. Tian, B. Chen, L. Dai, Z. Duan, and M. Zhou, "Deep variational graph convolutional recurrent network for multivariate time series anomaly detection," in *Int. Conf. Mach. Learn. (ICML)*. PMLR, 2022, pp. 3621–3633.
- [151] S. Han and S. S. Woo, "Learning sparse latent graph representations for anomaly detection in multivariate time series," in *Proc. 28th ACM SIGKDD Conf. Knowl. Discov. Data Min. (KDD)*, 2022, pp. 2977–2986.
- [152] Y. Tang, L. Zhao, S. Zhang, C. Gong, G. Li, and J. Yang, "Integrating prediction and reconstruction for anomaly detection," *Pattern Recognit. Lett.*, vol. 129, pp. 123–130, 2020.
- [153] M. Zheng, J. Man, D. Wang, Y. Chen, Q. Li, and Y. Liu, "Semi-supervised multivariate time series anomaly detection for wind turbines using generator scada data," *Reliab. Eng. Syst. Saf.*, vol. 235, p. 109235, 2023.
- [154] Y. Wei, J. Jang-Jaccard, W. Xu, F. Sabrina, S. Camtepe, and M. Boulic, "Lstm-autoencoder-based anomaly detection for indoor air quality time-series data," *IEEE Sens. J.*, vol. 23, no. 4, pp. 3787–3800, 2023.
- [155] G. Pu, L. Wang, J. Shen, and F. Dong, "A hybrid unsupervised clustering-based anomaly detection method," *Tsinghua Sci. Technol.*, vol. 26, no. 2, pp. 146–153, 2020.
- [156] B. Liu, Y. Xiao, L. Cao, Z. Hao, and F. Deng, "Svdd-based outlier detection on uncertain data," *Knowl. Inf. Syst.*, vol. 34, pp. 597–618, 2013.
- [157] A. P. Muniyandi, R. Rajeswari, and R. Rajaram, "Network anomaly detection by cascading k-means clustering and c4. 5 decision tree algorithm," *Procedia Eng.*, vol. 30, pp. 174–182, 2012.
- [158] A. M. Ikotun, A. E. Ezugwu, L. Abualigah, B. Abuhaiba, and J. Heming, "K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data," *Inf. Sci.*, vol. 622, pp. 178–210, 2023.
- [159] H. V. Singh, A. Girdhar, and S. Dahiya, "A literature survey based on dbscan algorithms," in *Proc. 2022 6th Int. Conf. Intell. Comput. Control Syst. (ICICCS)*. IEEE, 2022, pp. 751–758.
- [160] F. Murtagh and P. Contreras, "Algorithms for hierarchical clustering: an overview," *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.*, vol. 2, no. 1, pp. 86–97, 2012.
- [161] J. Li, H. Izakian, W. Pedrycz, and I. Jamal, "Clustering-based anomaly detection in multivariate time series data," *Appl. Soft Comput.*, vol. 100, p. 106919, 2021.
- [162] A. Markovitz, G. Sharir, I. Friedman, L. Zelnik-Manor, and S. Avidan, "Graph embedded pose clustering for anomaly detection," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2020, pp. 10 539–10 547.
- [163] S. Qiu, J. Ye, J. Zhao, L. He, L. Liu, E. Bicong, and X. Huang, "Video anomaly detection guided by clustering learning," *Pattern Recognit.*, vol. 153, p. 110550, 2024.

- [164] I. Kobyzev, S. J. Prince, and M. A. Brubaker, "Normalizing flows: An introduction and review of current methods," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 11, pp. 3964–3979, 2020.
- [165] J. Yu, Y. Zheng, X. Wang, W. Li, Y. Wu, R. Zhao, and L. Wu, "Fastflow: Unsupervised anomaly detection and localization via 2d normalizing flows," *arXiv preprint arXiv:2111.07677*, 2021.
- [166] M. Cho, T. Kim, W. J. Kim, S. Cho, and S. Lee, "Unsupervised video anomaly detection via normalizing flows with implicit latent features," *Pattern Recognit.*, vol. 129, p. 108703, 2022.
- [167] Q. Zhou, S. He, H. Liu, J. Chen, and W. Meng, "Label-free multivariate time series anomaly detection," *IEEE Trans. Knowl. Data Eng.*, 2024.
- [168] E. Dai and J. Chen, "Graph-augmented normalizing flows for anomaly detection of multiple time series," *arXiv preprint arXiv:2202.07857*, 2022.
- [169] Y. Zhou, X. Liang, W. Zhang, L. Zhang, and X. Song, "Vae-based deep svdd for anomaly detection," *Neurocomputing*, vol. 453, pp. 131–140, 2021.
- [170] Z. Zhang and X. Deng, "Anomaly detection using improved deep svdd model with data structure preservation," *Pattern Recognit. Lett.*, vol. 148, pp. 1–6, 2021.
- [171] J. Luo, J. Lin, Z. Yang, and H. Liu, "Smd anomaly detection: A self-supervised texture–structure anomaly detection framework," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–11, 2022.
- [172] C.-L. Li, K. Sohn, J. Yoon, and T. Pfister, "Cutpaste: Self-supervised learning for anomaly detection and localization," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2021, pp. 9664–9674.
- [173] R. Liu, W. Liu, Z. Zheng, L. Wang, L. Mao, Q. Qiu, and G. Ling, "Anomaly-gan: A data augmentation method for train surface anomaly detection," *Expert Syst. Appl.*, vol. 228, p. 120284, 2023.
- [174] Q. Wen, L. Sun, F. Yang, X. Song, J. Gao, X. Wang, and H. Xu, "Time series data augmentation for deep learning: A survey," *arXiv preprint arXiv:2002.12478*, 2020.
- [175] H. Hojjati, T. K. K. Ho, and N. Armanfard, "Self-supervised anomaly detection in computer vision and beyond: A survey and outlook," *Neural Networks*, vol. 172, p. 106106, 2024.
- [176] X. Zhang, M. Xu, and X. Zhou, "Realnet: A feature selection network with realistic synthetic anomaly for anomaly detection," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2024, pp. 16 699–16 708.
- [177] F. Van Wyk, Y. Wang, A. Khojandi, and N. Masoud, "Real-time sensor anomaly detection and identification in automated vehicles," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 3, pp. 1264–1276, 2019.
- [178] M. Abououf, R. Mizouni, S. Singh, H. Otrok, and E. Damiani, "Self-supervised online and lightweight anomaly and event detection for iot devices," *IEEE Internet Things J.*, vol. 9, no. 24, pp. 25 285–25 299, 2022.
- [179] X. Zhou, J. Wu, W. Liang, I. Kevin, K. Wang, Z. Yan, L. T. Yang, and Q. Jin, "Reconstructed graph neural network with knowledge distillation for lightweight anomaly detection," *IEEE Trans. Neural Netw. Learn. Syst.*, 2024.
- [180] Y. Zhao, G. H. Chen, and Z. Jia, "Tod: Gpu-accelerated outlier detection via tensor operations," *arXiv preprint arXiv:2110.14007*, 2021.
- [181] A. Al-Mazrawe and B. Al-Musawi, "Anomaly detection in cloud network: A review," in *BIO Web of Conferences*, vol. 97. EDP Sciences, 2024, p. 00019.
- [182] Z. Niu, G. Zhong, and H. Yu, "A review on the attention mechanism of deep learning," *Neurocomputing*, vol. 452, pp. 48–62, 2021.
- [183] H. Liu, X. Huang, M. Jia, T. Jia, J. Han, Y. Li, and Z. Wu, "Uac-ad: Unsupervised adversarial contrastive learning for anomaly detection on multi-modal data in microservice systems," *IEEE Trans. Serv. Comput.*, 2024.