

Review

Data Augmentation and Knowledge Transfer-Based Fault Detection and Diagnosis in Internet of Things-Based Solar Insecticidal Lamps: A Survey

Zhengjie Wang ¹, Xing Yang ^{1,*}, Tongjie Li ^{1,*}, Lei Shu ^{2,3,4,*}, Kailiang Li ⁴ and Xiaoyuan Jing ^{5,6,7}

¹ College of Intelligent Manufacturing, Anhui Science and Technology University, Chuzhou 233100, China; zhengjiewang1@outlook.com

² NAU-Lincoln Joint Research Center of Intelligent Engineering, Nanjing Agricultural University, Nanjing 210031, China

³ School of Engineering, University of Lincoln, Lincoln LN6 7TS, UK

⁴ College of Artificial Intelligence, Nanjing Agricultural University, Nanjing 210031, China; kailiang_li@njau.edu.cn

⁵ School of Computer, Guangdong University of Petrochemical Technology, Maoming 525000, China; xiaoyuanjing@whu.edu.cn

⁶ Guangdong Provincial Key Laboratory of Petrochemical Equipment Fault Diagnosis, Guangdong University of Petrochemical Technology, Maoming 525000, China

⁷ School of Computer Science, Wuhan University, Wuhan 430072, China

* Correspondence: xingyang@ahstu.edu.cn (X.Y.); litz@ahstu.edu.cn (T.L.); lei.shu@njau.edu.cn (L.S.)

Abstract

Internet of Things (IoT)-based solar insecticidal lamps (SIL-IoTs) offer an eco-friendly alternative by merging solar energy harvesting with intelligent sensing, advancing sustainable smart agriculture. However, SIL-IoTs encounter practical challenges, e.g., hardware aging, electromagnetic interference, and abnormal data patterns. Therefore, developing an effective fault detection and diagnosis (FDD) system is essential. In this survey, we systematically identify and address the core challenges of implementing FDD of SIL-IoTs. Firstly, the fuzzy boundaries of sample features lead to complex feature interactions that increase the difficulty of accurate FDD. Secondly, the category imbalance in the fault samples limits the generalizability of the FDD models. Thirdly, models trained on single scenarios struggle to adapt to diverse and dynamic field conditions. To overcome these challenges, we propose a multi-level solution by discussing and merging existing FDD methods: (1) a data augmentation strategy can be adopted to improve model performance on small-sample datasets; (2) federated learning (FL) can be employed to enhance adaptability to heterogeneous environments, while transfer learning (TL) addresses data scarcity; and (3) deep learning techniques can be used to reduce dependence on labeled data; these methods provide a robust framework for intelligent and adaptive FDD of SIL-IoTs, supporting long-term reliability of IoT devices in smart agriculture.

Keywords: agricultural pest control; Internet of Things-based solar insecticidal lamps; fault detection and diagnosis; data augmentation; knowledge transfer



Academic Editor: Jiangtao Huangfu

Received: 8 July 2025

Revised: 30 July 2025

Accepted: 4 August 2025

Published: 5 August 2025

Citation: Wang, Z.; Yang, X.; Li, T.; Shu, L.; Li, K.; Jing, X. Data Augmentation and Knowledge Transfer-Based Fault Detection and Diagnosis in Internet of Things-Based Solar Insecticidal Lamps: A Survey. *Electronics* **2025**, *14*, 3113. <https://doi.org/10.3390/electronics14153113>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The global population is increasing steadily, while the area of arable land is shrinking, making food security a critical issue that cannot be overlooked [1]. Currently, agricultural pests and diseases are one of the primary factors that damage plant growth and then threaten food-production security [2,3], leading to a yearly crop yield reduction of

20–40% [4]. Insecticides can effectively protect crops from agricultural pests and diseases. However, the overuse of insecticides poses significant risks to human health and damages the natural environment [5].

Solar insecticidal lamps (SIL) are an environmentally friendly pest-control technology that reduces pollution and effectively manages pest populations [6]. As shown in the Figure 1, the integration of SILs with Internet of Things (IoT) technology has led to the development of agricultural IoT devices, namely IoT-based solar insecticidal lamps (SIL-IoTs). These systems utilize smart sensors to acquire data related to agricultural pests, including species, population, and activity patterns, while attracting and eliminating agricultural pests. These data can be transmitted to the back-end or edge systems for analysis and processing, enabling farmers to monitor and manage agricultural pests [7].

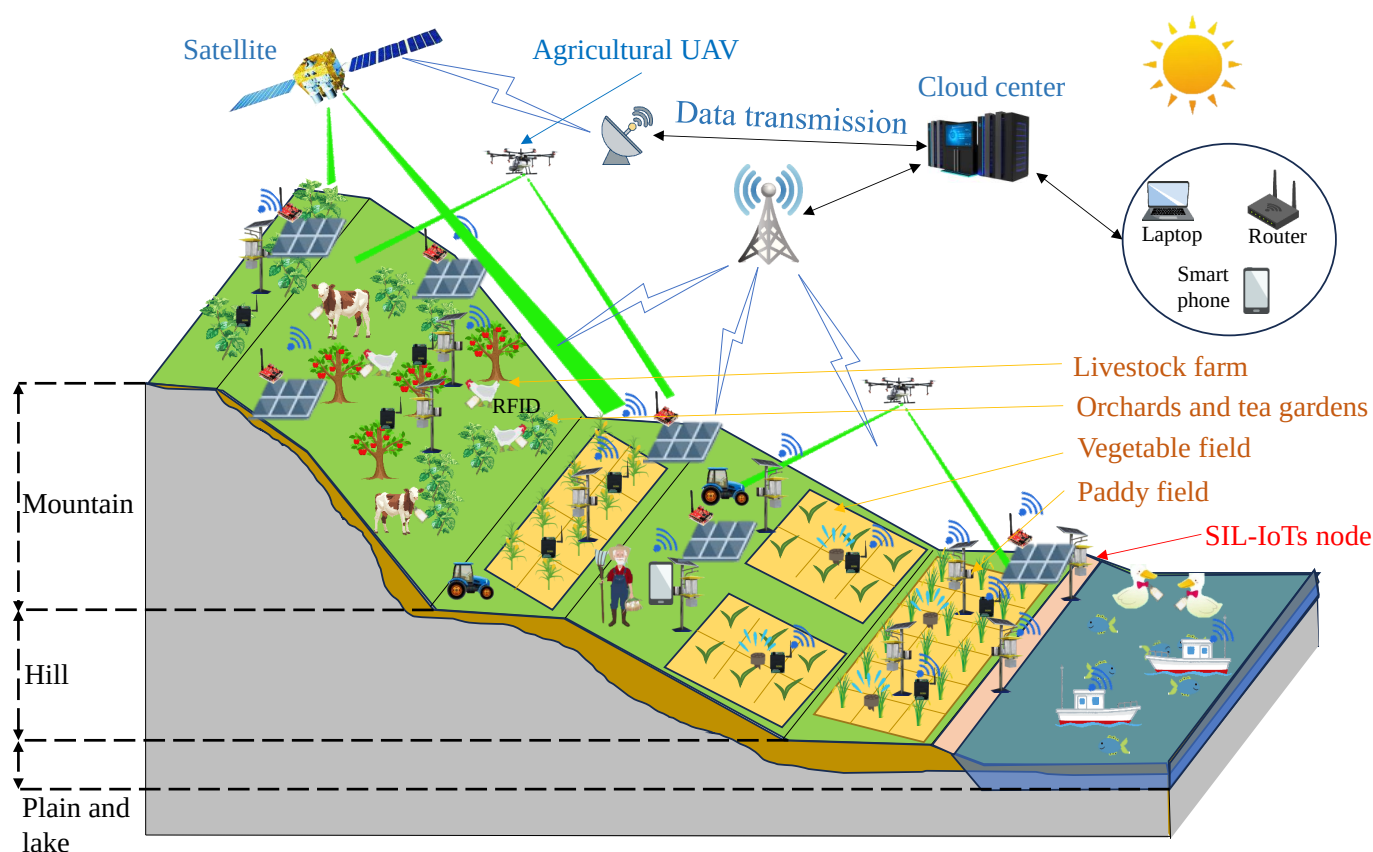


Figure 1. SIL-IoTs typical application scenarios: (1) livestock, (2) farmland, and (3) aquaculture. In addition, SIL-IoTs nodes are widely used in different terrains, e.g., mountains, hills and plains.

Additional sensing modules can be incorporated into SIL-IoTs to monitor agricultural parameters, e.g., weather conditions, soil moisture, and crop growth. Although agricultural weather networks already cover large regions, SIL-IoTs can be deployed in the field as nodes, providing micro-scale environmental monitoring and enhancing the multi-scale monitoring capabilities of existing agricultural weather networks [8]. Furthermore, SIL-IoTs can track soil moisture and acquire data on crop conditions, allowing for more accurate yield predictions [9]. Table 1 describes Acronyms.

To adapt diverse agricultural environments such as livestock, farmland, and aquaculture, the data collected by SIL-IoTs can also be transmitted to the cloud, allowing farmers to easily monitor and manage their operations remotely [10]. The extended functionalities of SIL-IoTs play a crucial role in advancing agricultural pest and disease management, contributing to increased agricultural productivity [11–14].

Table 1. Acronyms and descriptions.

Acronym	Description	Acronym	Description
IoT	Internet of Things	SIL	Solar insecticidal lamp
FDD	Fault detection and diagnosis	BSW	Binary sliding window-based
1DCNN	1D convolutional neural network	TL	Transfer learning
DA	Domain adaptation	GAN	Generative adversarial networks
VAE	Variational autoencoder	FEMA	Failure mode and effects analysis
FL	Federated learning	VFL	Vertical federated learning
FTL	Federated transfer learning	WSNs	Wireless sensor networks
SSL	Self-supervised learning	UL	Unsupervised learning
CWRU	Case western reserve university	LSTM	Long short-term memory network
AE	Auto-encoder	ML	Machine learning
RPN	Risk priority number	CSM-SSL	Cross-sensor multidimensional SSL
SLDDA	SSL classification adaptive mode	DGMM	Dual Gaussian mixture Model

Since SIL-IoTs are deployed in outdoor environments for extended periods, they are highly susceptible to environmental factors and prone to failure. Therefore, a robust fault detection and diagnosis (FDD) system is essential for timely fault detection and repair. SIL-IoTs use high-voltage metal mesh to eliminate pests. However, pest corpses often adhere to the metal mesh, which reduces the effectiveness of pest control [15]. In addition, SIL-IoTs are typically installed in elevated positions, leading to difficulties in maintenance and cleaning. Consequently, since most traditional SIL lack FDD capabilities, maintenance processes must manually detect faults. This significantly increases overall maintenance costs. By integrating FDD functionality, SIL-IoTs can autonomously detect and analyze faults in real time, thereby improving system reliability and extending its operational lifespan.

Based on this, we survey FDD methods for SIL-IoT. The results reveal that existing FDD methods primarily depend on manual inspections and empirical judgments, which not only delay FDD but also exhibit low accuracy. According to a review of intelligent FDD systems for SIL-IoTs [16], these systems offer real-time monitoring of SIL-IoTs operation, facilitating timely FDD through data analysis. Although current FDD methods for SIL-IoTs can accurately and timely detect faults to a certain extent, they still present limitations:

- Since SIL-IoTs are deployed in the wild for a long time and work in a harsh environment, devices are more prone to failure. Multiple faults occurring simultaneously can evolve into compound faults, affecting the accuracy of FDD. Therefore, the accuracy and precision of FDD methods need to be improved.
- SIL-IoTs mainly relies on solar energy for energy supply. Insufficient residual energy during continuous bad weather influences the normal operation of the device. Therefore, a lightweight design is required for FDD methods to save energy.
- Since changes in the device deployment environment will lead to a decrease in the generalization of the FDD method, updating the model will increase the cost of labeling the data. Therefore, data augmentation is needed to reduce the cost of labeling data.

Based on these, we conduct a review on FDD of SIL-IoTs methods, and the main contributions are:

- To investigate the potential issues of FDD of SIL-IoTs, we analyze and categorize the reasons and categories of fault generation.
- To address the current limitations of FDD of SIL-IoTs in practice, we analyze the following issues and point out the corresponding challenges, i.e., compound faults, sample labeling imbalance, single scenarios, computational and energy constraints.
- According to the challenges of FDD of SIL-IoTs, we identify the corresponding countermeasures including data augmentation and TL.

This paper is structured as follows: Section 2 provides an overview of the hardware and software components of SIL-IoTs and analyzes their fault characteristics based on recent relevant literature. Section 3 reviews the current methods used for FDD in SIL-IoTs and summarizes the limitations encountered in these scenarios. Section 4 outlines potential future research directions to address these limitations. Finally, Section 5 concludes the paper.

2. Characterization of SIL-IoTs

In this section, we introduce the hardware and software of SIL-IoTs and analyze their fault characteristics. We begin by presenting the hardware and software components along with their operational flow. Subsequently, we delve into the limitations and characteristics of SIL-IoTs related to faults.

2.1. Hardware and Software Introduction of SIL-IoTs

This section focuses on analyzing the hardware architecture of SIL-IoTs and their fault-generation mechanisms. Traditional SILs are generally classified into two types: light-trapping SIL, which uses high-voltage metal mesh to kill pest [17], and fan suction-based SIL, which relies on negative pressure airflow to capture and eliminate pests [18]. By comparing both SILs shown in Figure 2, it is evident that light-trapping SILs are more effective at killing larger pests, such as beetles and locusts, through the discharge of high-voltage arcs from the metal mesh. On the other hand, fan-suction SIL targets smaller pests, e.g., flies and mosquitoes, by inhaling them into the system [18]. Light-trapping SIL is susceptible to hardware degradation due to its reliance on high-frequency discharges. Sustained arcing can cause carbonization damage to the metal mesh, reducing insecticidal performance. The electromagnetic noise produced also distorts sensor signals, reducing the reliability of the data acquisition system [19,20].

Traditional SILs typically lack IoT modules however, SIL-IoTs integrate various components based on the Arduino platform, including modules for clock synchronization, sensors, and data transmission. This configuration enables the remote FDD, e.g., abnormal switching times, temperature mismatches between the interior and exterior of the electrical box, and sensor data anomalies [21]. Based on recent literature, the selection of components for FDD methods is summarized. Table 2 provides a detailed description of the component selection strategy, offering valuable references for subsequent FDD [22,23].

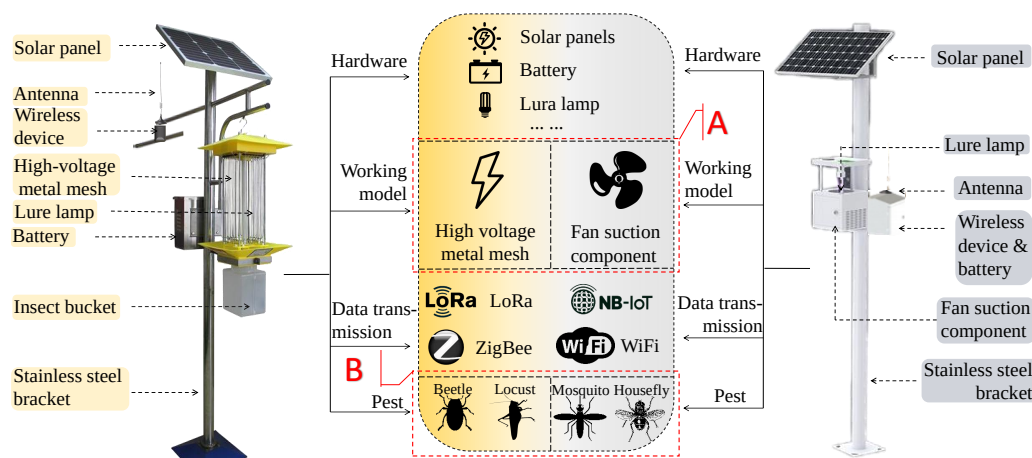


Figure 2. Difference between light-trapping and fan suction-based SIL. (A) Working mechanism comparison. (B) Targeted pest species comparison.

Table 2. Hardware components and their descriptions. According to the selection in the relevant literature, we describe the microcontroller, sensors, and working components.

Hardware	Description
Arduino (ATMEGA328PB)	Arduino can carry multiple sensor modules and read the data from the sensors. It has the advantage of high speed, low energy consumption, and cheap price [24].
Sensor module	1. DHT11 is a humidity and temperature sensor with NTC for the measurement of temperature and an 8-bit microcontroller for serial data output [25]. 2. MAX44009 is a light-intensity sensor that installs on solar panels and detects the intensity of sunlight radiation [26]. 3. DS18B20 is a waterproof temperature sensor; the temperature measurement range is -55 to $+125$ °C, and the operating voltage is 3 to 5.5 V [27].
High-voltage metal mesh Lure lamp	High-voltage metal mesh discharge 6 kV pulse to eliminate pests by electrical discharge [28]. 15 W LED lure lamp can effectively lure leaf borer, stem borer, and other types of rice lice [29].

The software structure and workflow of the SIL-IoTs system are illustrated in Figure 3. Firstly, the system harvests energy via solar panels and stores it in batteries. Next, pests are attracted by the lure lamps, and data related to the pests, e.g., species and activity patterns, are collected by smart sensors. Finally, the pests are eliminated through the high-voltage metal mesh, and the number of pests is recorded.

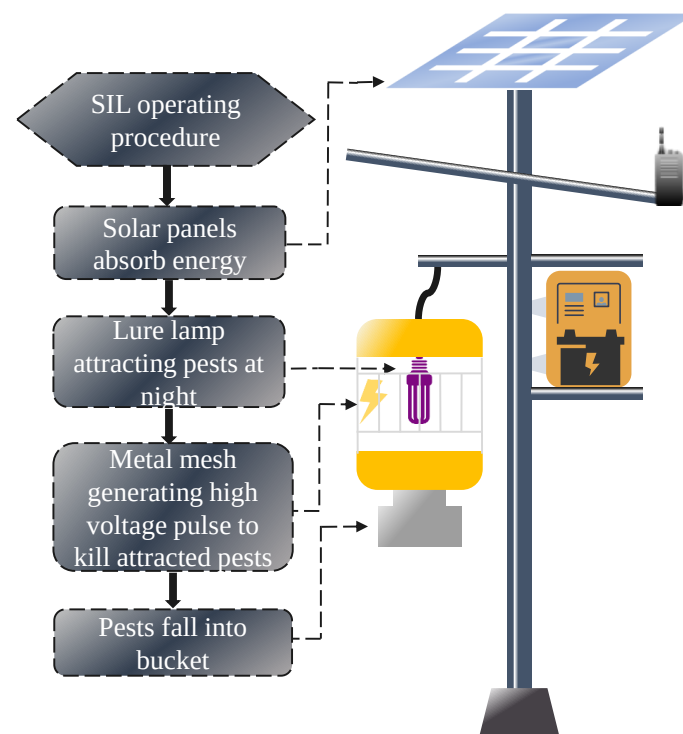


Figure 3. Working process of SIL-IoTs.

2.2. Characterization of SIL-IoTs

The accuracy and efficiency of FDD methods can only be ensured through a thorough analysis and classification of fault characteristics. Therefore, this section analyzes the deployment, data, and fault-label characteristics in detail.

- **Harsh deployment environment:** SIL-IoTs are usually deployed outdoors, where they operate in harsh environments and components are easy to be damaged and aged, resulting in frequent hardware and software failures. Hard faults are defined as faults where a component breaks down and causes the device to fail to work properly. For example, if the lure lamp is broken, SIL-IoTs cannot lure pests. Soft faults are defined as faults where the device is affected by an abnormal state of a

component, but is able to maintain operation. For example, sensor misalignment leading to data anomalies.

- **Low-quality data:** Data quality is a common challenge in IoT applications. As an IoT application scenario, SIL-IoTs requires high-quality data to ensure the accuracy of the FDD model [30]. However, due to the complexity of the environments in which these devices operate, the data often fails to meet expectations or becomes anomalous during transmission. Therefore, the data must be both informative and selective [31].
- **Few labeled samples:** In FDD, the number of labeled samples is usually limited, and samples often contain a large amount of data that does not match expectations [32]. This will lead to a decrease in the accuracy and generalization ability of the FDD model, and even trigger model overfitting or underfitting problems [33]. To solve this problem, data augmentation methods are widely used. Through data augmentation, not only can the number of samples be increased, but various fault characteristics can also be covered, thus ensuring that the FDD model can identify different types of faults.
- **Imbalance data category:** The FDD of SIL-IoTs needs to identify and analyze the causes of faults to ensure the normal operation of the devices. Thus, constructing effective FDD models requires a sufficient amount of labeled data for both training and evaluation [34]. In addition, to enhance the generalization capability of FDD of SIL-IoTs, it is necessary to label the newly generated data during device operation in different scenarios, allowing the models to adapt to various conditions. Therefore, SSL or TL can effectively solve the problem of labeled data imbalance.

2.3. Summaries

Based on the above analysis, it is necessary to consider the characteristics of SIL-IoTs when designing FDD methods for SIL-IoTs. Firstly, the quality of data plays a crucial role, with poor data quality significantly reducing the accuracy of FDD models. Secondly, the limited training samples hinder the knowledge transfer of these models to real-world scenarios. Finally, imbalanced data labeling complicates the model training process, leading to a decrease in the accuracy of the FDD model.

3. Related Method

This section summarizes methods applied to the FDD of SIL-IoTs based on relevant studies. Then, it analyzes the limitations of knowledge transfer-based FDD approaches for SIL-IoTs and presents solutions for similar scenarios.

3.1. FDD Methods Targeted to SIL-IoTs

Based on the classification and analysis of SIL-IoT fault characteristics shown in Figure 4, this section summarizes the current FDD of SIL-IoTs. Figure 4 shows an overview of various fault characteristics, such as label imbalance, device heterogeneity, and low data quality. These challenges will affect the performance of FDD models. Hardware and software failures, component differences, and environmental variations, among other factors, will lead sample feature bias, thereby worsening the aforementioned fault characteristics. This section discusses existing FDD methods, their limitations, and potential improvements.

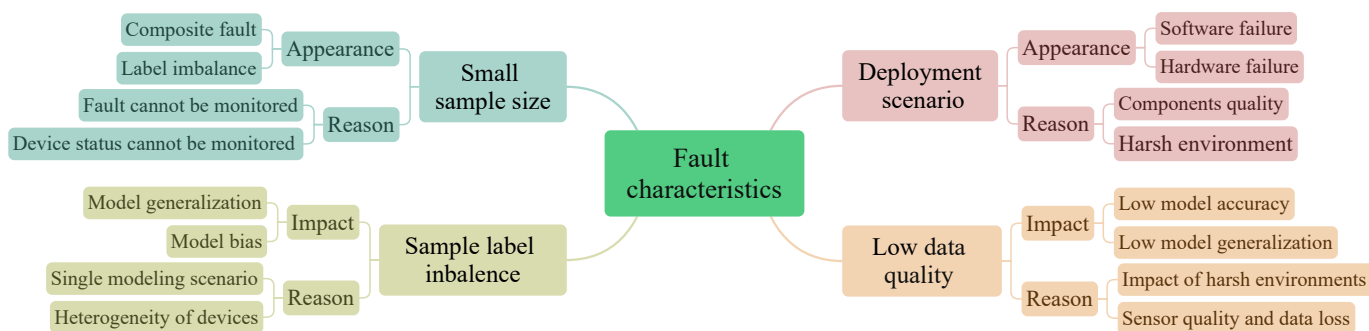


Figure 4. Fault characteristics appear in different conditions and cause different impacts. The left side is presented for the sample, while the right side is presented for the environment.

Binary sliding-window-based fault self-detection scheme (BSW) [35]: BSW is a lightweight, low false-detection rate, and low energy consumption self-detection method. The BSW method works as follows. Firstly, BSW constructs fault rules (e.g., fault code, fault importance) and associated fault phenomena (e.g., sensor data offset, data mismatch) using a fault dictionary. Then, it applies the binary sliding-window technique using eight binary digits to determine whether the current data aligns with the corresponding rules in the fault dictionary. Experimental results show that the BSW method achieves an average accuracy of 99.15% and reduces energy consumption by 71% during data transmission. The BSW method utilizes accuracy as the evaluation metric, as shown in the following formula:

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

where TP represents true positive (actual fault, predicted as fault); TN represents true negative (actual normal, predicted as normal); FP represents false positive (actual normal, predicted as fault); and FN represents false negative (actual fault, predicted as normal). However, some limitations persist regarding accuracy: (1) Due to its inherent storage limitations, the Arduino chip can only match fault information to entries within a predefined fault dictionary, thereby hindering its ability to precisely pinpoint the specific faulty module. Consequently, information from neighboring nodes becomes essential for refined FDD. (2) When constructing the FDD model using the fault dictionary, the fault rules are predefined. As a result, these original rules may not be applicable in different scenarios.

Based on the above issues, scholars have proposed a sensor-level lightweight fault-detection scheme [22]. This scheme handles faults undetectable by the BSW method, e.g., mismatches between solar panel current and light intensity, differences between air temperature and internal box temperature, and failures within the clock chip. This scheme diagnoses these faults by analyzing: (1) interval number residuals, which refer to the differences in interval numbers obtained by dividing the distribution of feature values (based on historical data) into intervals and then numbering them, and comparing the faulty node's interval number with its neighbors'. (2) Work anomaly differences, meaning the differences in features between nodes operating in different states. (3) Characteristic residuals, calculated by finding the difference in eigenvalues between the faulty node and its neighbors, and then determining the fault by accumulating the residuals. Experimental results demonstrate that this method achieves an F1-score of 92.42% and 95.59% for one-hop and two-hop nodes, respectively, while reducing data energy consumption for node information exchange by 25%. The sensor-level lightweight fault-detection scheme uses accuracy, F1-score, and Pearson correlation coefficient (r) as evaluation metrics. The formulas for F1-score and Pearson correlation coefficient (r) are given as follows:

$$F_1\text{-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (2)$$

where *Precision* is $\frac{TP}{TP+FP}$, *Recall* is $\frac{TP}{TP+FN}$.

$$r = \frac{\sum_{k=1}^n (x_{i,k} - \bar{X}_i) \cdot (x_{j,k} - \bar{X}_j)}{\sigma_i \cdot \sigma_j} \quad (3)$$

In this case, $\bar{X}_i$ and $\bar{X}_j$ are variable means. σ_i and σ_j are standard deviations. It is important to note that this method's accuracy is not absolute, as it can be affected by signal interference (e.g., sensor failures caused by high-voltage electromagnetic pulses).

The above methods are only suitable for situations where fault characteristics are more apparent. In cases where the sensor data characteristics are less obvious, particularly in sensor data drift, these methods may not be effective. To address this, related researchers have proposed a lightweight separable 1D convolutional neural network (SA1D-CNN) that can be implemented in SIL-IoTs nodes to identify sensor faults, reduce detection delay, and minimize data transmission [36]. The SA1D-CNN utilizes accuracy, F1-score, and FLOPs as evaluation metrics. The formula for FLOPs is given as follows:

$$\text{FLOPs}_{\text{sep}} = k \times 1 \times C \times W \times 1 + C \times C' \times W \times 1 \quad (4)$$

where k is kernel size, C is number of input channels, C' is number of output channels, W is input feature size. Experimental results show that this method achieves an accuracy of 99.9%, with an average F1-score of 97.6%, while consuming only 4.33 W of energy and 351 KB of memory. However, a limitation of this method is that it struggles to prioritize faults, making it difficult to detect severe faults in the early stages. Since there are limited FDD methods for SIL-IoTs, we investigate FDD methods for similar scenarios.

Figure 5 provides an overview of the current FDD methods in SIL-IoTs scenarios. The figure mainly organizes three approaches along several critical dimensions:

- The BSW method is suitable for addressing critical faults, such as operation unit failures and energy system malfunctions. Sensor-level schemes are typically used to diagnose routine faults, such as equipment operating in abnormal states, whereas SA1DCNN is commonly employed to identify subtle faults, including sensor data inconsistencies.
- Lightweight self-diagnostic techniques such as BSW and SA1DCNN eliminate the need for data from neighboring nodes. In contrast, sensor-level schemes typically involve analyzing the data weights of neighboring nodes to perform fault diagnosis.
- BSW relies on fault dictionaries and binary pattern matching. A sensor-level scheme analyzes working states and calculates feature residuals. The SA1DCNN, a deep learning method, utilizes depthwise separable convolutions and attention mechanisms.

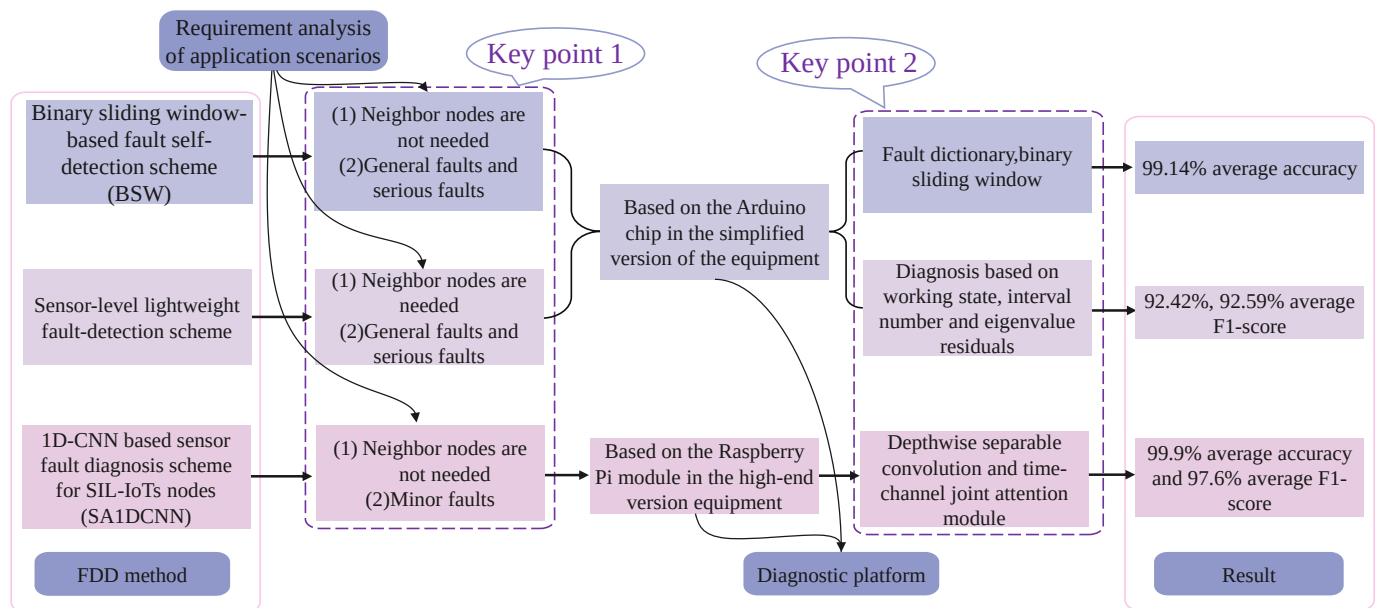


Figure 5. Current FDD methods described. Key points are (1) whether neighbor node information is required and (2) what type of faults can be identified.

3.2. Data Augmentation-Based FDD Schemes

Fault labels are crucial for FDD model training, yet compound faults are challenging to acquire and label, limiting their practical use [30]. SIL-IoTs face diverse faults influenced by environmental factors, whereas regular faults can be labeled for supervised learning, compound faults often evade accurate labeling, hindering conventional FDD analysis. Data quality, affected by sensor performance and acquisition methods, critically impacts FDD accuracy and efficiency. Poor data may reduce model performance or prolong training, necessitating preprocessing or augmentation techniques.

Scholars have utilized data augmentation methods to significantly improve data sensitivity and specificity across various domains including image recognition, acoustic features, and text data [37,38]. For high-dimensional signals like acoustic waves, data augmentation improves model robustness by enhancing time and frequency-domain signals through noise injection or adversarial networks. For text data, it reduces labeling costs by randomly inserting and deleting elements. Studies show augmented data can boost model accuracy by up to 20% versus original datasets [39,40]. Figure 6 shows different data augmentation strategies, such as SMOTE, generative model-based methods, and data transformation based methods. Generative models like GANs and VAEs, for instance, can prove beneficial in tasks involving few-shot learning or zero-shot learning [41]. These techniques are particularly valuable for SIL-IoTs applications, addressing challenges like data quality and compound faults. Data augmentation enhances domain robustness and generates higher-quality data across source and target domains [42]. Data augmentation also helps mitigate sample labeling imbalance, a prevalent challenge in FDD stemming from (1) insufficient fault-type diversity and (2) feature distortion due to transmission issues. These imbalances pose a risk of majority-class overfitting and compromised FDD sensitivity. Such issues, however, can be effectively addressed through self-supervised learning (SSL) and unsupervised learning (UL) approaches.

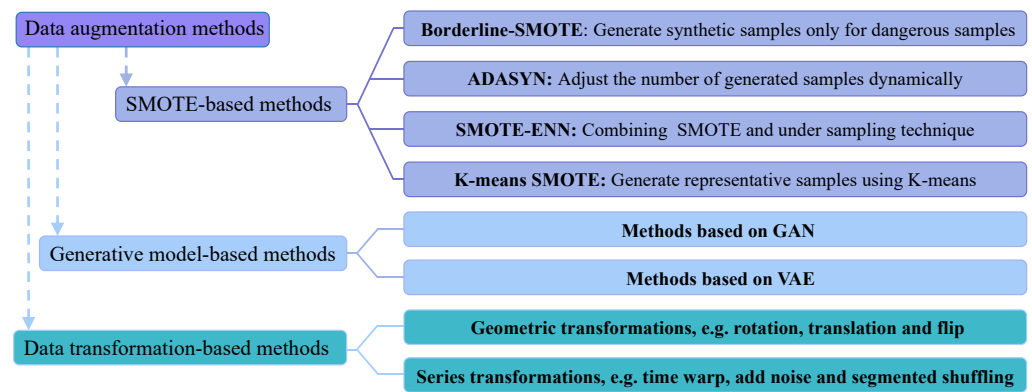


Figure 6. Classifications of data augmentation methods.

To address the aforementioned issues, we explored solutions from similar scenarios, e.g., rotating machinery, IoT devices, and other industrial equipment. SSL enhances the ability of model to recognize minority-class samples by leveraging latent information in unlabeled data. It can efficiently train features from large-scale unlabeled datasets using context-based training [43]. Scholars have applied SSL for FDD in various scenarios through contrastive [44–48] and generative [49–53] approaches, exceeding an accuracy of 90%. These SSL methods not only effectively identify multiple fault types but also reduce the reliance on labeled data. This feature makes SSL highly flexible and adaptable, especially in SIL-IoTs, where diverse compound faults and uneven sample labeling are common. However, SSL requires a large number of unlabeled samples for training, while SIL-IoTs typically involve small-sample FDD. To overcome this limitation, UL approaches can be adopted to leverage existing domain knowledge and a small number of labeled samples, thereby enhancing the performance of FDD models.

To deal with high-dimensional data and reduce noise interference and feature distortion, researchers in this field have adopted unsupervised learning (UL) to address noise and distortion problems in datasets by analyzing data features and relationships across both time and frequency domains [54]. In deeper applications, these methods are able to significantly reduce the dimensionality while still retaining important information in the data, thus enhancing the robustness and recognition ability of the model [55]. It is shown that UL models are able to effectively use patterns and information in unlabeled data to provide accurate FDD even without relying on labeled samples [56]. Studies on rotating-machinery fault diagnosis have demonstrated that features extracted from different dimensions (time domain, frequency domain and time–frequency domain) can significantly reduce the interference of dissimilar features. These UL-based approaches are equally applicable to FDD of SIL-IoTs. Due to challenges in SIL-IoTs such as sample label imbalance, low model-generalization capability, and difficulties in knowledge-transfer processes, filtering the feature data becomes particularly important as it helps to reduce model training time while maintaining accuracy. Research shows that models can achieve 90% accuracy even when trained with only 40 fault-sample labels [57]. Given that distortion often present in SIL-IoT fault data can easily lead to model training overfitting, the UL strategy demonstrates high feasibility for such applications. However, UL may introduce model bias towards majority classes during training. Therefore, it needs to be combined with knowledge transfer techniques to ensure the stability and accuracy of the model in dealing with complex scenarios.

3.3. Knowledge Transfer-Based FDD Schemes

FDD models for SIL-IoTs often suffer from limited generalizability due to single-scenario training. Transfer learning addresses this by enabling cross-domain knowledge

transfer [58]. Variations in environmental factors such as light intensity, electromagnetic interference, or device heterogeneity between deployment and training scenarios can cause shifts in sensor data distribution. This misalignment often leads to a significant decline in FDD accuracy and, in extreme cases, can result in complete functional failure. Thus, data augmentation enhances model generalization by increasing data samples, particularly beneficial when deal with low-quality target domain data [58].

To address sensor data drift from device heterogeneity, researchers employ transfer learning (TL) and domain adaptation (DA) techniques to optimize FDD models. TL effectively overcomes domain generalization challenges by transferring knowledge from source to target domains, accelerating training while mitigating data drift and reducing resource consumption. As illustrated in the Figure 7, TL methods typically involve extracting features from the source domain and then leveraging dimensional differences to identify and separate fault features, thus capturing the distinction between the source and target domains [59,60].

To address data bias and class imbalance in multi-modal or cross-sensor fusion scenarios, Dai et al. [61] demonstrated that utilizing a multi-modal cross-sensor transformer can substantially enhance the generalization capabilities across different sensors in various FDD tasks. Lin et al. [62] addressed the sensor sample imbalance problem by optimizing the FDD model through a multi-sensor cross-fusion strategy. By employing a multidimensional modeling strategy for sensor signals, Zhu et al. [63] were able to both capture the correlated features of different sensors and narrow the feature gap between individual channels. Li et al. [64] studied the distribution bias in sensor data through domain adaptive techniques and thereby achieved differentiated transfer of FDD models between distinct domains. To address the issue of inconsistent labels across sensor data, Xu et al. [65] employed time–frequency analysis to identify similar data within time–frequency signals, ultimately achieving cross-domain knowledge transfer.

For example, in wind-turbine FDD, fault accuracy using dimensionally extracted fault data has achieved up to 99% [66]. Addressing the limited adaptability of traditional FDD models in diverse deployment scenarios, DA improves their performance in novel scenarios by utilizing feature difference analysis for model transfer [67]. By optimizing classification loss, central judgment loss, and alignment loss between domains, DA ensures intra-class compactness and inter-class separability while maintaining feature consistency, demonstrating 94.8% accuracy on the CWRU bearing dataset [68]. Together, TL and DA address complementary aspects of cross-domain FDD optimization.

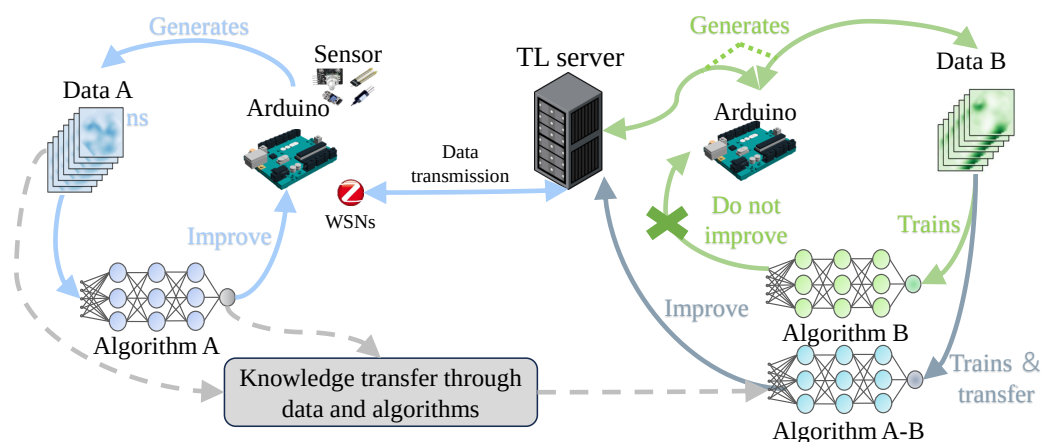


Figure 7. In the process of TL in FDD, a model trained on fault data from domain A can be applied to domain B, as there are some shared characteristics between the two domains. As a result, a new model (A–B) can be generated by integrating data from domain B through TL.

As shown in the Figure 8, firstly, we discuss the current categorization of federated learning (FL). Specifically, it can be divided into vertical federated learning, horizontal federated learning, and federated transfer learning. Then, we analyze the heterogeneity in terms of samples and features, e.g., same features but different sample, same sample but different features, and different sample and different features. Finally, data privacy and data security issues of different FL methods are considered. Furthermore, appropriate WSN components should be selected based on practical needs. For challenges like data silos and sensitivity, FL enables collaborative analysis while operating locally on data, minimizing leakage risks and addressing privacy concerns [68]. Researchers have explored FL in FDD for rotating machinery [69–71] and IoT devices [72–74]. These studies show that FL effectively addresses privacy and security concerns while maintaining 99% accuracy and improving model robustness. In IoT applications, FL reduces edge-cloud gradient disparity, significantly lowering computational costs for SIL-IoTs FDD implementations. Moreover, FL enables collaborative learning across nodes to enhance overall model capability [75].

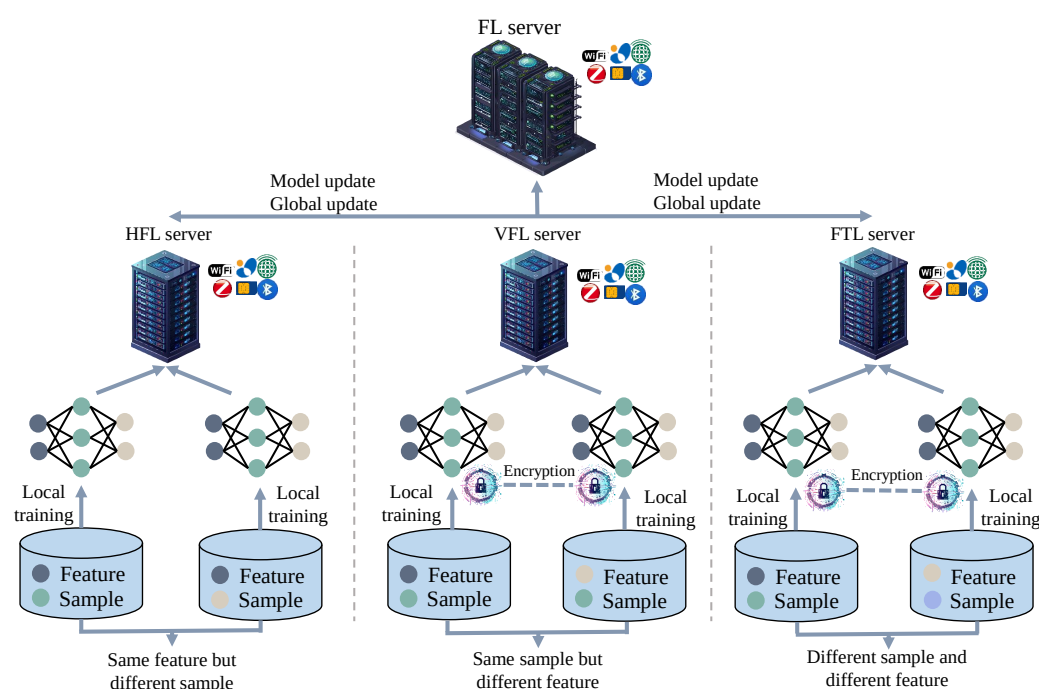


Figure 8. Categorization of FL, and the judgment of whether encryption is needed based on models trained on different samples and features.

3.4. Failure Mode and Effects Analysis-Based FDD Schemes

Failure mode and effects analysis (FMEA) can assess the importance of fault data to improve model performance, such as by evaluating data availability, optimizing data augmentation strategies, and enhancing the model's reliability under actual operating conditions [76]. In addition, FMEA analyzes fault modes to determine which knowledge are worth transfer while conducting a comprehensive risk assessment [77]. For example, fault data is generated in the spatial and temporal dimensions (e.g., frequency, duration, and location of failures) that may not be apparent through static analysis alone. By using this dynamic information, FMEA can prioritize fault modes that pose the highest risk to system performance, safety, or reliability. This allows for more targeted and effective transfer strategies and ensures that resources are allocated to address the most critical vulnerabilities in the system [78].

FMEA serves as a critical risk assessment tool across industries, systematically identifying failure modes and their system impacts [79]. For SIL-IoTs in harsh environments,

FMEA enables accurate failure classification along three dimensions: behavioral (interfering and non-interfering), temporal (transient, intermittent and permanent), and spatial (global and node-specific) [16]. This dimensional approach dynamically prioritizes faults while reducing FDD computational complexity through optimal resource allocation. Research demonstrates FMEA's effectiveness in embedded systems, where it reduces decision tree operations and sensor communication overhead [80]. The method evaluates fault severity, incidence, detectability, and diagnostic priority using weighting factors [80,81], achieving notable diagnostic accuracy for industrial faults: 82.97% (temperature), 79.78% (voltage), and 77% (current) [82].

FMEA is also well-suited for SIL-IoTs, as these systems rely on Arduino microcontrollers with limited computational resources. This strategy helps to effectively minimize the waste of computing resources and enables earlier detection of component faults in harsh operating conditions. Moreover, it allows for a deep analysis of the fault modes, facilitating accurate identification of the root causes. Additionally, by prioritizing faults based on importance assessment and screening those with a more sufficient sample size, this approach helps balance sample labels, thereby improving both model training efficiency and accuracy. Based on the FDD approach for SIL-IoTs discussed in the previous section and the associated fault characteristics, this section presents FDD schemes for similar scenarios, tailored to the corresponding characteristics. The advantages and limitations of these approaches are summarized in Table 3.

3.5. The Limitations of the Above Methods

The above methods have some challenges when adopted in SIL-IoTs:

- Data augmentation and low label dependency related FDD methods require large amounts of high-quality data. However, the datasets for FDD of SIL-IoTs suffer from significant interference and distortion, necessitating the use of appropriate data-preprocessing techniques.
- Although FEMA can effectively evaluate the importance of faults, it usually does not consider differences in deployment environments and does not adjust for data heterogeneity, making it difficult to respond accurately to complex fault information.
- Due to the deployment of SIL-IoTs in various scenarios, there can be differences in terrain, crop types, and target pests. The components of SIL-IoTs in different scenarios will also be different, thus the data between SIL-IoTs nodes varies. For these reasons, applying FL and TL to SIL-IoTs presents challenges, as the data heterogeneity and varying system configurations can affect the accuracy and effectiveness of these methods.

Table 3. Related studies and their characteristics.

Method	Author	Application Scenarios	F1-Score	Accuracy	Advantage	Disadvantage
Data Augmentation	Zeng et al. [39]	• Smart agriculture	N/A	92.6%	• Improve generalizability	• High computing cost
	Abayomi-Alli et al. [40]	• Industrial IoTs	96.76%	97.7%	• Increases sample size	• Ineffective sample augment
	Kim et al. [83]	• Geographic information systems	99%	100%	• Suitable for different tasks	• Data inconsistency
	Stivaktakis et al. [84]		77.7%	85.7%		
Self-Supervised Learning	Wang et al. [50]	• Gearbox	N/A	97.32%	• Reduces label dependency	• High computing cost
	Wan et al. [52]	• Bearings	97.85%	98.41%	• Enhances adaptability	• Low generalization
	Wei et al. [53]	• Industrial IoTs	N/A	93.4%	• Improves robustness	• Data sensitivity
	Yang et al. [69]		87.21%	99.98%	• Improve feature learning	• Requires unlabeled data
Unsupervised Learning	Brito et al. [54]	• Railway systems	92.78%	N/A	• Reduces labeling cost	• No clear objectives
	Guo et al. [55]	• Gearbox	96.67%	99%	• Handles insufficient data	• High complexity
	Wang et al. [56]	• Industry IoTs	N/A	99.5%	• Extensive applications	• Lack of evaluation
Federated Learning	Yang et al. [46]	• Industrial IoTs	N/A	100%	• Reduces annotation dependency	• High data quality needs
	Shen et al. [70]	• Bearings	N/A	98.47%	• Reduces training time	• Negative transfer
	Geng et al. [71]	• Power system	87.76%	95.56%	• Enhances generalization	• High computing cost
Transfer Learning	Qian et al. [66]	• Power system	N/A	98%	• Reduces computation	• High-quality data needed
	He et al. [85]	• Geographic systems	99%	100%	• Reduces data needs	• Model negative transfer
	Shao et al. [86]	• Rotating machinery	N/A	98.8%	• Enhances generalization	• Domain mismatch
			RPN _{min}	RPN _{max}		
Failure Mode and Effects Analysis	Kadena et al. [81]	• Economic field	24.3	300	• Risk assessment	• Experience dependent
	Tarcsay et al. [82]	• Industrial IoTs	0	120	• Process improvement	• Incomplete risk coverage
	Ma et al. [87]		2.584	22.128		

In risk analysis, RPN is utilized to quantify risk levels; RPN_{min} represents the lowest risk priority, while RPN_{max} indicates the highest. The effectiveness of models or systems can be evaluated through measures like accuracy (the ratio of correct predictions) and the F1-score (the harmonic mean of precision and recall).

4. Challenges and Future Directions

This section scrutinizes the challenges confronting existing FDD methods within SIL-IoT scenarios and investigates prospective solutions. We begin by analyzing the limitations of current approaches, considering the unique characteristics and constraints of SIL-IoTs datasets. Subsequently, optimization strategies are explored through a comparative analysis with datasets and solutions from other domains. Concluding, we outline future research directions, guided by current research progress, to foster advancements in SIL-IoT FDD.

4.1. Analysis the Challenges of Similar Datasets

The correlation FDD methods are confronted with multiple challenges when applied to SIL-IoTs, mainly in four aspects, i.e., low data quality, scarcity of compound fault samples, lack of model generalization, and limited resources at the edge end. These challenges not only affect the effectiveness and applicability of FDD methods, but also limit its generalization in practical applications. To gain a deep understanding of these issues, the reasons for the emergence of these challenges and their impact on the effectiveness of the methods are analyzed in the context of specific dataset characteristics.

Through an investigation into the existing FDD of SIL-IoTs, several key issues have been identified. For instance, current approaches often focus solely on sensor data abnormalities or hardware circuit faults, with limited attention given to compound faults. Additionally, there is a scarcity of fault samples involving compound faults. The fault data often contains a high level of noise and low quality, traditional data augmentation techniques struggle to effectively capture fault features. Over-augmentation can even lead to the loss of crucial features. Furthermore, data augmentation tends to increase energy consumption, which further limits the scalability of the FDD model.

Another challenge is that SIL-IoTs are easy to occur novel fault types based on their deployment environment. These new faults can fall outside the scope of predefined fault rules. For example, the existing BSW method [35] is heavily reliant on a fault dictionary, making it difficult to dynamically adapt to complex fault combinations. In comparison, datasets from similar scenarios, such as the CWRU bearing dataset [88], exhibit high FDD accuracy with rich fault signals and controlled experimental setups. These datasets are typically well-structured and feature distinct fault signatures, making them suitable for training FDD models under ideal conditions. However, they lack a simulation of dynamic environmental disturbances, such as variations in temperature, humidity, or interference, which are common in real-world applications, e.g., SIL-IoTs. For example, while the FDD model trained on the CWRU dataset can achieve accuracy greater than 96%, it struggles with handling issues such as sensor data drift caused by sudden weather changes in SIL-IoTs [89].

In the context of FDD of SIL-IoTs, the scarcity of compound fault-sample labels has been identified as a major constraint on model generalization performance [22,35,36]. This issue is particularly pronounced due to the disproportionate number of normal samples compared to faulty ones, with certain fault classes being underrepresented. An additional challenge arises from feature distortion, which is aggravated by packet loss in WSNs, leading to model bias towards multi-class overfitting and diminishing the ability to accurately detect critical faults. Although UL and SSL techniques can mitigate the reliance on labeled data, UL struggles to distinguish between normal and abnormal states when fault data is absent. Meanwhile, SSL requires substantial volumes of unlabeled data for effective training in small-sample scenarios, e.g., SIL-IoTs. Notably, datasets such as the Intel Lab Data dataset [90], which collects multidimensional environmental data (e.g., light, temperature, and humidity) from 54 sensors, offer a valuable source of training samples for

both UL and SSL approaches. Furthermore, its structured feature-extraction methodology provides useful insights for advancing cross-domain FDD strategies in SIL-IoTs.

Table 4 lists typical datasets for smart agriculture, e.g., Smart Farming Data 2024 (SF24) [91] and Agriculture and Farming Dataset [92]. SIL-IoTs are susceptible to cross-scene distribution bias in their sensor data, as they are typically deployed as nodes in agricultural fields. Indeed, models trained in a single scene often experience a significant drop in accuracy when transferred to new environments due to domain shift. This phenomenon occurs because the model has learned patterns specific to the training environment, which may not hold in different scenarios with varying environmental conditions, sensor characteristics, or fault patterns. While TL and DA can mitigate some of the domain differences, device heterogeneity, e.g., hardware batch variations and aging of components will result in negative migration, necessitating the design of more robust feature-alignment strategies. The problem is further exacerbated in agricultural IoT deployments, where the spatio-temporal variability of the environment amplifies data heterogeneity. This leads to gradient conflicts and convergence stability challenges during global model aggregation in traditional FL. Moreover, while FL offers performance improvements through distributed collaboration, its deployment at the edge faces several bottlenecks. Firstly, narrow-band communication protocols, e.g., ZigBee, with latencies between 0.1–0.3 s [93], hinder high-frequency gradient synchronization, potentially increasing fault leakage rates by up to 41% [91]. Moreover, the latency associated with inference of complex FDD model makes it challenging to meet real-time demands. Additionally, the long-term field deployment of equipment, subject to hardware wear and tear, e.g., pest adhesion on metal mesh [16] further hampers the efficiency of the pest-control function and leads to high maintenance costs [91], underscoring the limitations of the related FDD methods in adapting to these challenges.

Table 4. Comparison of agricultural IoT datasets and their recorded parameters.

Dataset Name	Group/ Sample Size	Temperature & Humidity	Soil Moisture	Sunlight Exposure	Wind Speed & Direction	Component Voltage	Component Current
FDD of SIL-IoTs dataset (https://ieee-dataport.org/documents/silf-dataset-faultdataset-solar-insecticidal-lamp-internet-things-node) [94] Accessed on 27 July 2025	16/469,916	✓	×	✓	✓	✓	✓
SIL-IoTs insecticidal count dataset (https://ieee-dataport.org/documents/insecticidal-counting-dataset-based-one-solar-insecticidal-lamp-and-two-cameras) [95] Accessed on 27 June 2025	6/708,480	×	×	✓	✓	✓	✓
Smart Farming Data 2024 (SF24) (https://www.kaggle.com/datasets/datasetengineer/smart-farming-data-2024-sf24) [91] Accessed on 18 April 2025	19/2200	✓	✓	✓	✓	×	×
Agriculture and Farming Dataset (https://www.kaggle.com/datasets/bhadramohit/agriculture-and-farming-dataset) [92] Accessed on 18 April 2025	9/50	✓	✓	✓	×	×	×

Table 4. Cont.

Dataset Name	Group/ Sample Size	Temperature & Humidity	Soil Moisture	Sunlight Exposure	Wind Speed & Direction	Component Voltage	Component Current
Irrigation machine dataset (https://www.kaggle.com/datasets/mahmoudshaheen1134/irrigation-machine-dataset) [96] Accessed on 18 April 2025	22/2000	✓	✓	✓	×	✓	✓
Intrusion detection systems dataset [97]	9/2,540,044	✓	×	×	×	✓	✓

4.2. Analysis the Challenges of Relevant Methods

When faced with compound fault features, traditional data augmentation methods often result in only a modest 20% increase in training data [38]. Excessive transformations, e.g., flipping, scaling, and shifting, may not only fail to capture essential fault characteristics but also generate invalid synthetic data, leading to the unnecessary consumption of computational resources [38]. To address this issue, recent research has proposed a joint architecture combining Long Short-Term Memory Networks (LSTMs) and Auto-Encoder (AE) algorithms. By leveraging the time-series modeling capability of LSTM to capture fault evolution patterns and the sparse feature characterization of AE for decoupling, this approach has demonstrated an improvement in the F1-score of bearing FDD, reaching up to 97.3% [83]. However, compound faults are prevalent in practical applications. Models trained on historical single faults often experience a higher false-alarm rate when exposed to novel compound faults, primarily due to imbalanced sample distribution. To overcome this issue, recent studies have utilized weighted oversampling techniques, such as SMOTE-Tomek, to balance small classes of samples, along with feature space adaptation for domain alignment. Experimental results indicate that this method achieves an accuracy of over 93% [98].

FMEA effectively assesses the importance of faults. However, traditional FMEA faces significant limitations in practical applications, e.g., human subjective bias, challenges in modeling compound faults, and sensitivity to data variations [32]. To address these shortcomings, recent studies have proposed integrating machine learning (ML) techniques, e.g., decision trees and random forests, into the FMEA framework. These integrations optimize the computational logic of the Risk Priority Number (RPN) by incorporating feature importance ranking. Experimental results demonstrate that this approach enhances the accuracy of compound fault identification to 91.4% [99]. Additionally, Bayesian networks, through probabilistic reasoning, can dynamically assess the impact of feature changes on the model, leading to a 52% reduction in fault under-reporting rates [100]. Notably, the potential of FMEA extends beyond risk assessment. It can also be leveraged to simulate assumptions and evaluate how changes in system characteristics during the FDD process influence model accuracy [101,102]. Moreover, the application of FMEA span data-driven decision-making and the development of FDD modeling frameworks [103].

ML approaches, e.g., SSL, UL, and FL, have been widely used in FD of industrial and IoT devices. However, large amounts of data are required, and category imbalance sensitivity and high-dimensional feature redundancy are still the main challenges for SSL. Relevant scholars have tried to solve the overfitting problem caused by category imbalance. A FDD method is proposed to extract features from a large number of unlabeled samples and construct FDD models [104]. For FDD of rotating machinery, it is found that the category imbalance will lead to the problem of model bias to the number of multi-classes. To solve this problem, a self-supervised learning classification adaptive model (SLDDA) is

presented, and the experimental results show that F1-score of SLDDA reaches 93.7% [105]. However, the features extracted from sensors are limited, and fault features usually need to be extracted from multiple sensors when compound faults occur. To balance such fault features, related scholars proposed a cross-sensor multidimensional self-supervised learning approach (CSM-SSL). The compound FDD accuracy is achieved 89.4% by fusing multi-signal time–frequency features and introducing an attention mechanism to optimize the feature weight allocation [106]. When dealing with deployment scenario differences and compound faults, UL can learn from some unlabeled data. However, when optimizing the sample features, due to the lack of normal data points around the faulty data points, making it difficult for the model to learn the difference between the normal state and the faulty state. Moreover, the lack of fault data in the model can lead to recognize faults incorrectly. Therefore, we compare the two approaches in Figure 9. Relevant scholars extracted fault features through the time dimension and then used a Dual Gaussian Mixture Model (DGMM) for FDD, and the experimental results showed that the FDD accuracy could reach 81.3% under the no-labeling condition [107]. In addition, the problem of model generalization decay due to data heterogeneity of agricultural IoT nodes can be effectively reduced by a FTL framework. As shown in Figure 9, we can clearly see the differences between SSL and UL in terms of label usage, model-generation methods, and downstream tasks.

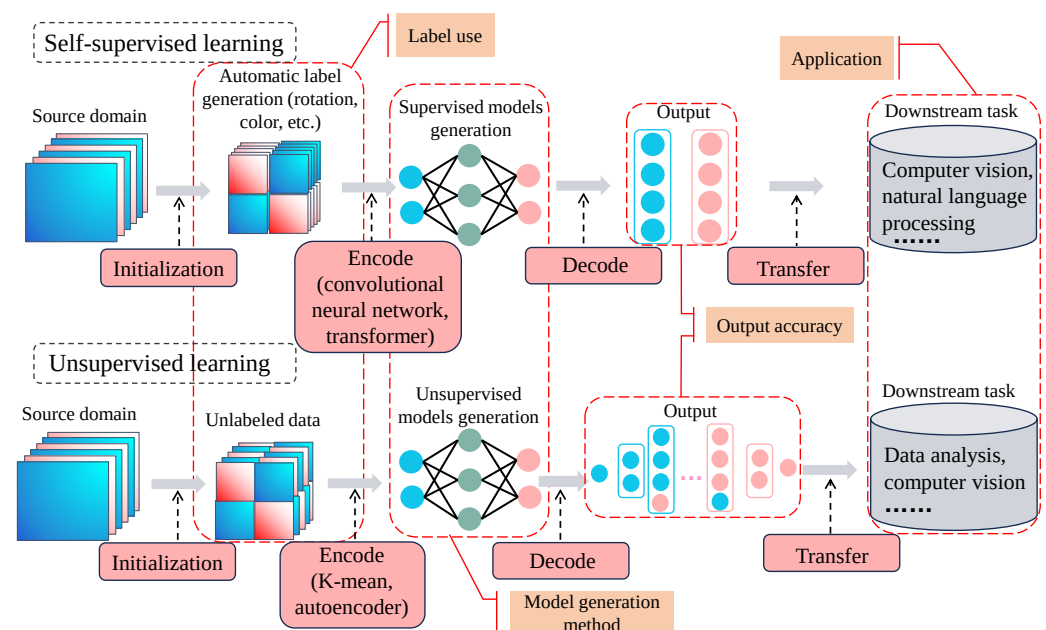


Figure 9. SSL and UL have some differences in label use, model generation, and downstream applications. Moreover, it can easily lead to model overfitting since UL does not need to use sample labels.

Although FL and TL have great application prospects in distributed FDD, they still face challenges, e.g., data heterogeneity, communication bottlenecks and aggregation conflicts. To cope with the above challenges, scholars have recently proposed a combination of FL and TL based on aligning the feature space of source and target domains by domain-invariant knowledge distillation, which has been shown to reduce the model generalization error in heterogeneous data scenarios to 8.3% [108]. In terms of communication overhead, some studies have used techniques, e.g., compression modeling and gradient sparsification, to reduce the amount of data transferred between the edge device and the cloud aggregator, resulting in a 78% reduced amount of communication data [38]. Relevant scholars determine the aggregation weights of the model on the basis of the similarity and

importance of the features. Then, the operations of feature extraction and transformation are performed on the model data from different IoT nodes, which improves the model convergence stability by 62% [109].

5. Conclusions

SIL-IoTs play a crucial role in green prevention and control of plant diseases and pests. FDD of SIL-IoTs is an important means to ensure that SIL-IoTs work reliably. We introduce the components and fault characteristics of SIL-IoTs, identifies the limitations of current FDD methods, and discusses FDD approaches (e.g., data augmentation, and knowledge transfer) in similar scenarios to enhance FDD efficiency of SIL-IoTs. While existing methods struggle with issues (e.g., low data quality, sample imbalance, and environmental variability), deep learning techniques (e.g., TL, SSL, and FL) offer promising solutions. These methods enable more adaptive, data-efficient, and resource-aware FDD solutions, especially in edge computing scenarios. By focusing on data augmentation and knowledge transfer, this research provides a comprehensive foundation for enhancing the FDD performance of SIL-IoTs and offers valuable insights for broader applications in smart agriculture.

Author Contributions: Z.W.: Investigation, Visualization, Writing—original draft. X.Y.: Conceptualization, Funding acquisition, Investigation, Writing—review and editing. T.L.: Funding acquisition, Writing—review and editing. L.S.: Supervision, Writing—review and editing. K.L.: Writing—review and editing. X.J.: Writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: The authors declare financial support was received for the research, authorship, and/or publication of this article. This research was funded in part by the National Natural Science Foundation of China under Grant 62402003, in part by the Talent Introduction Project under Grant ZNYJ202402.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors would like to thank the reviewers for their valuable and detailed comments, which were essential to improve this report. Their valuable and detailed comments were crucial in improving the quality of this paper.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Qiu, B.; Li, H.; Tang, Z.; Chen, C.; Berry, J. How cropland losses shaped by unbalanced urbanization process? *Land Use Policy* **2020**, *96*, 104715. [\[CrossRef\]](#)
2. Luneja, R.L.; Mkindi, A.G. Advances in Botanical-Based Nanoformulations for Sustainable Cotton Insect Pest Management in Developing Countries. *Front. Agron.* **2024**, *7*, 1558395. [\[CrossRef\]](#)
3. Qian, Y.; Xiao, Z.; Deng, Z. Fine-grained Crop Pest Classification based on Multi-scale Feature Fusion and Mixed Attention Mechanisms. *Front. Plant Sci.* **2025**, *16*, 1500571. [\[CrossRef\]](#)
4. Dhanda, S.; Singh, N.; Ikley, J.T.; DeWerff, R.; Werle, R.; Sarangi, D. Dicamba-Based Preemergence Herbicide Tank Mixtures Improved Residual Weed Control in Dicamba-Resistant Soybean. *Front. Agron.* **2024**, *7*, 1576547. [\[CrossRef\]](#)
5. Lykogianni, M.; Bempelou, E.; Karamaouna, F.; Aliferis, K.A. Do pesticides promote or hinder sustainability in agriculture? The challenge of sustainable use of pesticides in modern agriculture. *Sci. Total Environ.* **2021**, *795*, 148625. [\[CrossRef\]](#)
6. Yang, F.; Shu, L.; Huang, K.; Li, K.; Han, G.; Liu, Y. A Partition-Based Node Deployment Strategy in Solar Insecticidal Lamps Internet of Things. *IEEE Internet Things J.* **2020**, *7*, 11223–11237. [\[CrossRef\]](#)
7. Abdollahi, A.; Rejeb, K.; Rejeb, A.; Mostafa, M.M.; Zailani, S. Wireless Sensor Networks in Agriculture: Insights from Bibliometric Analysis. *Sustainability* **2021**, *13*, 12011. [\[CrossRef\]](#)

8. Kelley, J.; McCauley, D.; Alexander, G.A.; Gray, W.F.; Siegfried, R.; Oldroyd, H.J. Using machine learning to integrate on-farm sensors and agro-meteorology networks into site-specific decision support. *Trans. ASABE* **2020**, *63*, 1427–1439. [\[CrossRef\]](#)
9. Singh, A.K.; Balabaygloo, B.J.; Bekee, B.; Blair, S.W.; Fey, S.; Fotouhi, F.; Gupta, A.; Jha, A.; Martinez-Palomares, J.C.; Menke, K.; et al. Smart connected farms and networked farmers to improve crop production, sustainability and profitability. *Front. Agron.* **2024**, *6*, 1410829. [\[CrossRef\]](#)
10. Ge, H.; Zhang, Q.; Shen, M.; Qin, Y.; Wang, L.; Yuan, C. Enhancing yield prediction in maize breeding using UAV-derived RGB imagery: A novel classification-integrated regression approach. *Front. Plant Sci.* **2025**, *16*, 1511871. [\[CrossRef\]](#) [\[PubMed\]](#)
11. Lloret, J.; Sendra, S.; Garcia, L.; Jimenez, J.M. A wireless sensor network deployment for soil moisture monitoring in precision agriculture. *Sensors* **2021**, *21*, 7243. [\[CrossRef\]](#) [\[PubMed\]](#)
12. Sharma, S.; Rai, S.; Krishnan, N.C. Wheat Crop Yield Prediction Using Deep LSTM Model. *arXiv* **2020**, arXiv:2011.01498. [\[CrossRef\]](#)
13. Puntel, L.A.; Thompson, L.J.; Mieno, T. Leveraging digital agriculture for on-farm testing of technologies. *Front. Agron.* **2024**, *6*, 1234232. [\[CrossRef\]](#)
14. Feng, J.; Blair, S.W.; Ayanlade, T.T.; Balu, A.; Ganapathysubramanian, B.; Singh, A.; Sarkar, S.; Singh, A.K. Robust soybean seed yield estimation using high-throughput ground robot videos. *Front. Plant Sci.* **2025**, *16*, 1554193. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Li, Y.; Du, B.; Luo, L.; Luo, Y.; Yang, X.; Liu, Y.; Shu, L. A Scheme for Pest-Dense Area Localization with Solar Insecticidal Lamps Internet of Things Under Asymmetric Links. *IEEE Trans. AgriFood Electron.* **2023**, *1*, 71–85. [\[CrossRef\]](#)
16. Yang, X.; Shu, L.; Huang, K.; Li, K.L.; Huo, Z.Q.; Wang, Y.F.; Wang, X.Y.; Lu, Q.L.; Zhang, Y.C. Characteristics analysis and challenges for fault diagnosis in solar insecticidal lamps internet of things. *Smart Agric.* **2020**, *2*, 11. (In Chinese with English Abstract) [\[CrossRef\]](#)
17. Li, K.L.; Shu, L.; Huang, K.; Sun, Y.H.; Yang, F.; Zhang, Y.; Huo, Z.Q.; Wang, Y.F.; Wang, X.Y.; Lu, Q.L.; et al. Research and prospect of solar insecticidal lamps internet of things. *Smart Agric.* **2019**, *1*, 13. (In Chinese with English Abstract) [\[CrossRef\]](#)
18. Han, Y.; Song, Z.; Yi, W.; Zhan, C. Design of and Experimentation with a Suction-Based Pest-Capture Machine for the Tea Pest *Empoasca vitis*. *Agriculture* **2024**, *14*, 964. [\[CrossRef\]](#)
19. Balamurugan, R.; Kandasamy, P. Effectiveness of portable solar-powered light-emitting diode insect trap: Experimental investigation in a groundnut field. *J. Asia-Pac. Entomol.* **2021**, *24*, 1024–1032. [\[CrossRef\]](#)
20. Huang, K.; Li, K.; Shu, L.; Yang, X.; Gordon, T.; Wang, X. High Voltage Discharge Exhibits Severe Effect on ZigBee-Based Device in Solar Insecticidal Lamps Internet of Things. *IEEE Wirel. Commun.* **2020**, *27*, 140–145. [\[CrossRef\]](#)
21. Zhu, T.; Cheng, X.; Li, C.; Li, Y.; Pan, C.; Lu, G. Decoding plant thermosensors: Mechanism of temperature perception and stress adaption. *Front. Plant Sci.* **2025**, *16*, 1560204. [\[CrossRef\]](#)
22. Yang, X.; Shu, L.; Li, K.; Nurellari, E.; Huo, Z.; Zhang, Y. A Lightweight Fault-Detection Scheme for Resource-Constrained Solar Insecticidal Lamp IoTs. *Sensors* **2023**, *23*, 6672. [\[CrossRef\]](#)
23. Yang, F.; Shu, L. A Trajectory-Inspired Node Deployment Strategy in Solar Insecticidal Lamps Internet of Things Under Coverage and Maintenance Cost Considerations. *IEEE Trans. AgriFood Electron.* **2024**, *2*, 28–42. [\[CrossRef\]](#)
24. Ichwana.; Nasution, I.S.; Sundari, S.; Rifky, N. Data Acquisition of Multiple Sensors in Greenhouse Using Arduino Platform. *Iop Conf. Ser. Earth Environ. Sci.* **2020**, *515*, 012011. [\[CrossRef\]](#)
25. Srivastava, D.; Kesarwani, A.; Dubey, S. Measurement of Temperature and Humidity by using Arduino Tool and DHT11. *Int. Res. J. Eng. Technol. (IRJET)* **2018**, *5*, 876–878.
26. Atika, Z.; Leow, W.; Iszaidy, I.; Irwan, Y.; Safwati, I.; Irwanto, M.; Wafi, N.; Saw, S. Development A Portable Solar Energy Measurement System. *J. Phys. Conf. Ser.* **2021**, *1962*, 012049. [\[CrossRef\]](#)
27. Baehaqi, M.; Rosyid, A.; Siswanto, A.; Subiyanta, E. Performance Testing of DHT11 and DS18B20 Sensors as Server Room Temperature Sensors. *Mestro J. Tek. Mesin Dan Elektro* **2023**, *5*, 6–11. [\[CrossRef\]](#)
28. Chen, Y.D.; Wang, J.C.; Zhang, J.H.; Cao, G.Y. Light source for comfortable lighting and trapping pests in tea gardens based on solar-like lighting. *Appl. Opt.* **2020**, *59*, 8459–8464. [\[CrossRef\]](#)
29. Tu, H.; Tang, N.; Hu, X.; Yao, Z.; Wang, G.; Wei, H. LED multispectral circulation solar insecticidal lamp application in rice field. *Trans. Chin. Soc. Agric. Eng.* **2016**, *32*, 193–197.
30. Challa, H.; Niu, N.; Johnson, R. Faulty Requirements Made Valuable: On the Role of Data Quality in Deep Learning. In Proceedings of the 2020 IEEE Seventh International Workshop on Artificial Intelligence for Requirements Engineering (AIRE), Zurich, Switzerland, 1 September 2020; pp. 61–69. [\[CrossRef\]](#)
31. Serna M., E.; Bachiller S., O.; Serna A., A. Knowledge meaning and management in requirements engineering. *Int. J. Inf. Manag.* **2017**, *37*, 155–161. [\[CrossRef\]](#)
32. Zou, L.; Li, Y.; Xu, F. An adversarial denoising convolutional neural network for fault diagnosis of rotating machinery under noisy environment and limited sample size case. *Neurocomputing* **2020**, *407*, 105–120. [\[CrossRef\]](#)

33. Pothuganti, S. Review on over-fitting and under-fitting problems in Machine Learning and solutions. *Int. J. Adv. Res. Electr. Electron. Instrum. Eng.* **2018**, *7*, 3692–3695. [\[CrossRef\]](#)
34. Bagora, P.; Ebrahimzadeh, A.; Wuhib, F.; Glitho, R.H. Data Labeling for Fault Detection in Cloud: A Test Suite-Based Active Learning Approach. In Proceedings of the 2023 IEEE 9th International Conference on Network Softwarization (NetSoft), Madrid, Spain, 19–23 June 2023; pp. 262–266. [\[CrossRef\]](#)
35. Yang, X.; Shu, L.; Li, K.; Huo, Z.; Shu, S.; Nurellari, E. SILOS: An Intelligent Fault Detection Scheme for Solar Insecticidal Lamp IoT with Improved Energy Efficiency. *IEEE Internet Things J.* **2023**, *10*, 920–939. [\[CrossRef\]](#)
36. Yang, X.; Shu, L.; Li, K.; Huo, Z.; Zhang, Y. SA1D-CNN: A Separable and Attention Based Lightweight Sensor Fault Diagnosis Method for Solar Insecticidal Lamp Internet of Things. *IEEE Open J. Ind. Electron. Soc.* **2022**, *3*, 291–303. [\[CrossRef\]](#)
37. Chen, H.; Chen, J.; Ding, J. Data Evaluation and Enhancement for Quality Improvement of Machine Learning. *IEEE Trans. Reliab.* **2021**, *70*, 831–847. [\[CrossRef\]](#)
38. Chen, S.; Paul, K.C.; Zhao, T. Enhancing Arc Fault Detection Performance through Data Augmentation with Artificial Intelligence Technology: An Approach to Time Series Dataset Enlargement. In Proceedings of the 2024 IEEE Energy Conversion Congress and Exposition (ECCE), Phoenix, AZ, USA, 20–24 October 2024; pp. 1866–1872. [\[CrossRef\]](#)
39. Zeng, Q.; Ma, X.; Cheng, B.; Zhou, E.; Pang, W. GANs-Based Data Augmentation for Citrus Disease Severity Detection Using Deep Learning. *IEEE Access* **2020**, *8*, 172882–172891. [\[CrossRef\]](#)
40. Abayomi-Alli, O.O.; Damaševičius, R.; Misra, S.; Maskeliūnas, R. Cassava disease recognition from low-quality images using enhanced data augmentation model and deep learning. *Expert Syst.* **2021**, *38*, e12746. [\[CrossRef\]](#)
41. McDonald, T.; Butler, S.W. Progress and Current Topics of JEDEC JC-70.1 Power GaN Device Quality and Reliability Standards Activity: Or: What is the Avalanche capability of your GaN Transistor? In Proceedings of the 2021 IEEE International Reliability Physics Symposium (IRPS), Monterey, CA, USA, 21–25 March 2021; pp. 1–6. [\[CrossRef\]](#)
42. Ng, N.; Cho, K.; Ghassemi, M. SSMB: Self-Supervised Manifold Based Data Augmentation for Improving Out-of-Domain Robustness. *arXiv* **2020**, arXiv:2009.10195.
43. Liu, Y.; Jin, M.; Pan, S.; Zhou, C.; Zheng, Y.; Xia, F.; Yu, P.S. Graph self-supervised learning: A survey. *IEEE Trans. Knowl. Data Eng.* **2023**, *35*, 5879–5900. [\[CrossRef\]](#)
44. Ding, Y.; Zhuang, J.; Ding, P.; Jia, M. Self-supervised pretraining via contrast learning for intelligent incipient fault detection of bearings. *Reliab. Eng. Syst. Saf.* **2022**, *218*, 108126. [\[CrossRef\]](#)
45. Liu, Y.; Wen, W.; Bai, Y.; Meng, Q. Self-supervised feature extraction via time–frequency contrast for intelligent fault diagnosis of rotating machinery. *Measurement* **2023**, *210*, 112551. [\[CrossRef\]](#)
46. Yang, Z.; Huang, Y.; Nazeer, F.; Zi, Y.; Valentino, G.; Li, C.; Long, J.; Huang, H. A novel fault detection method for rotating machinery based on self-supervised contrastive representations. *Comput. Ind.* **2023**, *147*, 103878. [\[CrossRef\]](#)
47. Zhang, W.; Chen, D.; Xiao, Y.; Yin, H. Semi-Supervised Contrast Learning Based on Multiscale Attention and Multitarget Contrast Learning for Bearing Fault Diagnosis. *IEEE Trans. Ind. Inform.* **2023**, *19*, 10056–10068. [\[CrossRef\]](#)
48. Chaitanya, K.; Erdil, E.; Karani, N.; Konukoglu, E. Contrastive learning of global and local features for medical image segmentation with limited annotations. *Adv. Neural Inf. Process. Syst.* **2020**, *33*, 12546–12558.
49. Sun, S.; Wang, T.; Yang, H.; Chu, F. Condition monitoring of wind turbine blades based on self-supervised health representation learning: A conducive technique to effective and reliable utilization of wind energy. *Appl. Energy* **2022**, *313*, 118882. [\[CrossRef\]](#)
50. Wang, H.; Liu, Z.; Ge, Y.; Peng, D. Self-supervised signal representation learning for machinery fault diagnosis under limited annotation data. *Knowl.-Based Syst.* **2022**, *239*, 107978. [\[CrossRef\]](#)
51. Fu, D.; Liu, J.; Zhong, H.; Zhang, X.; Zhang, F. A novel self-supervised representation learning framework based on time-frequency alignment and interaction for mechanical fault diagnosis. *Knowl.-Based Syst.* **2024**, *295*, 111846. [\[CrossRef\]](#)
52. Wan, W.; Chen, J.; Zhou, Z.; Shi, Z. Self-Supervised Simple Siamese Framework for Fault Diagnosis of Rotating Machinery with Unlabeled Samples. *IEEE Trans. Neural Netw. Learn. Syst.* **2024**, *35*, 6380–6392. [\[CrossRef\]](#)
53. Wei, M.; Liu, Y.; Zhang, T.; Wang, Z.; Zhu, J. Fault Diagnosis of Rotating Machinery Based on Improved Self-Supervised Learning Method and Very Few Labeled Samples. *Sensors* **2022**, *22*, 192. [\[CrossRef\]](#) [\[PubMed\]](#)
54. Brito, L.C.; Susto, G.A.; Brito, J.N.; Duarte, M.A.V. Fault Detection of Bearing: An Unsupervised Machine Learning Approach Exploiting Feature Extraction and Dimensionality Reduction. *Informatics* **2021**, *8*, 85. [\[CrossRef\]](#)
55. Guo, Z.; Wan, Y.; Ye, H. An Unsupervised Fault-Detection Method for Railway Turnouts. *IEEE Trans. Instrum. Meas.* **2020**, *69*, 8881–8901. [\[CrossRef\]](#)
56. Wang, Y.; Zhou, J.; Zheng, L.; Gogu, C. An end-to-end fault diagnostics method based on convolutional neural network for rotating machinery with multiple case studies. *J. Intell. Manuf.* **2022**, *33*, 809–830. [\[CrossRef\]](#)
57. Brito, L.C.; Susto, G.A.; Brito, J.N.; Duarte, M.A. An explainable artificial intelligence approach for unsupervised fault detection and diagnosis in rotating machinery. *Mech. Syst. Signal Process.* **2022**, *163*, 108105. [\[CrossRef\]](#)
58. Su, J.; Yu, X.; Wang, X.; Wang, Z.; Chao, G. Enhanced transfer learning with data augmentation. *Eng. Appl. Artif. Intell.* **2024**, *129*, 107602. [\[CrossRef\]](#)

59. Wang, Y.; Yan, J.; Yang, Z.; Zhang, W.; Wang, J.; Geng, Y.; Srinivasan, D. A Class Alignment Multisource Domain Adaptation for Partial Discharge Condition Assessment with Unknown Faults in GIS. *IEEE Internet Things J.* **2025**, *12*, 19955–19971. [\[CrossRef\]](#)
60. Wang, Q.; Taal, C.; Fink, O. Integrating Expert Knowledge with Domain Adaptation for Unsupervised Fault Diagnosis. *IEEE Trans. Instrum. Meas.* **2022**, *71*, 3500312. [\[CrossRef\]](#)
61. Dai, Y.; Li, J.; Mei, Z.; Ni, Y.; Guo, S.; Li, Z. Self-Supervised Learning for Multimodal Fault Diagnosis with Shapley-Value Weighted Transformers. *IEEE Trans. Instrum. Meas.* **2025**, *74*, 3534114. [\[CrossRef\]](#)
62. Lin, T.; Ren, Z.; Huang, K.; Zhang, X.; Zhu, Y.; Yan, K.; Hong, J. Contribution Imbalance and the Improvement Method in Multisensor Information Fusion-Based Intelligent Fault Diagnosis of Rotating Machinery. *IEEE Trans. Instrum. Meas.* **2025**, *74*, 3525614. [\[CrossRef\]](#)
63. Zhu, H.; Zhao, Y.; Yan, X.; Kang, Y.; Liu, B. Cross-Sensor Generative Self-Supervised Learning Network for Fault Detection Under Few Samples. *J. Syst. Sci. Complex.* **2025**, *38*, 1000–1020. [\[CrossRef\]](#)
64. Li, X.; Deng, W.; Ncube, S.N.; Ming, R.; Duan, C.; Qin, Y.; Liu, F.; Luo, J.; Pu, H. Multi-Source Joint Adaptive Distribution with Online Transfer Learning for Cross-Domain Fault Diagnosis. *IEEE Trans. Reliab.* **2025**, 1–13. [\[CrossRef\]](#)
65. Xu, X.; Yang, X.; He, C.; Shi, P.; Hua, C. Adversarial Domain Adaptation Model Based on LDTW for Extreme Partial Transfer Fault Diagnosis of Rotating Machines. *IEEE Trans. Instrum. Meas.* **2024**, *73*, 3538811. [\[CrossRef\]](#)
66. Qian, Q.; Wu, F.; Wang, Y.; Qin, Y. Maximum subspace transferability discriminant analysis: A new cross-domain similarity measure for wind-turbine fault transfer diagnosis. *Comput. Ind.* **2025**, *164*, 104194. [\[CrossRef\]](#)
67. Ma, P.; Zhang, H.; Fan, W.; Wang, C. A diagnosis framework based on domain adaptation for bearing fault diagnosis across diverse domains. *ISA Trans.* **2020**, *99*, 465–478. [\[CrossRef\]](#)
68. Li, L.; Fan, Y.; Tse, M.; Lin, K.Y. A review of applications in federated learning. *Comput. Ind. Eng.* **2020**, *149*, 106854. [\[CrossRef\]](#)
69. Yang, D.; Long, Z.; Ma, X.; Xu, B. Wind Power Fault Detection Based on Vertical Federated Learning. In Proceedings of the 2023 Global Reliability and Prognostics and Health Management Conference (PHM-Hangzhou), Hangzhou, China, 12–15 October 2023; pp. 1–7. [\[CrossRef\]](#)
70. Shen, J.; Yang, S.; Zhao, C.; Ren, X.; Zhao, P.; Yang, Y.; Han, Q.; Wu, S. FedLED: Label-Free Equipment Fault Diagnosis with Vertical Federated Transfer Learning. *IEEE Trans. Instrum. Meas.* **2024**, *73*, 3509910. [\[CrossRef\]](#)
71. Geng, D.; He, H.; Lan, X.; Liu, C. Bearing fault diagnosis based on improved federated learning algorithm. *Computing* **2022**, *104*, 1–19. [\[CrossRef\]](#)
72. Zhang, W.; Lu, Q.; Yu, Q.; Li, Z.; Liu, Y.; Lo, S.K.; Chen, S.; Xu, X.; Zhu, L. Blockchain-Based Federated Learning for Device Failure Detection in Industrial IoT. *IEEE Internet Things J.* **2021**, *8*, 5926–5937. [\[CrossRef\]](#)
73. Liu, Y.; Garg, S.; Nie, J.; Zhang, Y.; Xiong, Z.; Kang, J.; Hossain, M.S. Deep Anomaly Detection for Time-Series Data in Industrial IoT: A Communication-Efficient On-Device Federated Learning Approach. *IEEE Internet Things J.* **2021**, *8*, 6348–6358. [\[CrossRef\]](#)
74. Li, Y.; Chen, Y.; Zhu, K.; Bai, C.; Zhang, J. An Effective Federated Learning Verification Strategy and Its Applications for Fault Diagnosis in Industrial IoT Systems. *IEEE Internet Things J.* **2022**, *9*, 16835–16849. [\[CrossRef\]](#)
75. Banabilah, S.; Aloqaily, M.; Alsayed, E.; Malik, N.; Jararweh, Y. Federated learning review: Fundamentals, enabling technologies, and future applications. *Inf. Process. Manag.* **2022**, *59*, 103061. [\[CrossRef\]](#)
76. Wang, L.; Tang, D.; Liu, C.; Nie, Q.; Wang, Z.; Zhang, L. An Augmented Reality-Assisted Prognostics and Health Management System Based on Deep Learning for IoT-Enabled Manufacturing. *Sensors* **2022**, *22*, 6472. [\[CrossRef\]](#) [\[PubMed\]](#)
77. Filz, M.A.; Langner, J.E.B.; Herrmann, C.; Thiede, S. Data-driven failure mode and effect analysis (FMEA) to enhance maintenance planning. *Comput. Ind.* **2021**, *129*, 103451. [\[CrossRef\]](#)
78. Amato, F.; Cirillo, E.; Fonisto, M.; Moccardi, A. Detecting Adversarial Attacks in IoT-Enabled Predictive Maintenance with Time-Series Data Augmentation. *Information* **2024**, *15*, 740. [\[CrossRef\]](#)
79. Liu, H.C.; Liu, L.; Liu, N. Risk evaluation approaches in failure mode and effects analysis: A literature review. *Expert Syst. Appl.* **2013**, *40*, 828–838. [\[CrossRef\]](#)
80. Li, Z.; Wang, Y.; Wang, K. A deep learning driven method for fault classification and degradation assessment in mechanical equipment. *Comput. Ind.* **2019**, *104*, 1–10. [\[CrossRef\]](#)
81. Kadena, E.; Koçak, S.; Takács-György, K.; Keszthelyi, A. FMEA in Smartphones: A Fuzzy Approach. *Mathematics* **2022**, *10*, 513. [\[CrossRef\]](#)
82. Rastayesh, S.; Bahrebar, S.; Blaabjerg, F.; Zhou, D.; Wang, H.; Dalsgaard Sørensen, J. A system engineering approach using FMEA and Bayesian network for risk analysis—A case study. *Sustainability* **2019**, *12*, 77. [\[CrossRef\]](#)
83. Kim, D.Y.; Kareem, A.B.; Domingo, D.; Shin, B.C.; Hur, J.W. Advanced Data Augmentation Techniques for Enhanced Fault Diagnosis in Industrial Centrifugal Pumps. *J. Sens. Actuator Netw.* **2024**, *13*, 60. [\[CrossRef\]](#)
84. Stivaktakis, R.; Tsagkatakis, G.; Tsakalides, P. Deep Learning for Multilabel Land Cover Scene Categorization Using Data Augmentation. *IEEE Geosci. Remote Sens. Lett.* **2019**, *16*, 1031–1035. [\[CrossRef\]](#)
85. He, Z.; Shao, H.; Zhong, X.; Zhao, X. Ensemble transfer CNNs driven by multi-channel signals for fault diagnosis of rotating machinery cross working conditions. *Knowl.-Based Syst.* **2020**, *207*, 106396. [\[CrossRef\]](#)

86. Shao, S.; McAleer, S.; Yan, R.; Baldi, P. Highly Accurate Machine Fault Diagnosis Using Deep Transfer Learning. *IEEE Trans. Ind. Inform.* **2019**, *15*, 2446–2455. [\[CrossRef\]](#)
87. Shao, J.; Zhong, S.; Tian, M.; Liu, Y. Combining fuzzy MCDM with Kano model and FMEA: A novel 3-phase MCDM method for reliable assessment. *Ann. Oper. Res.* **2024**, *342*, 725–765. [\[CrossRef\]](#)
88. Neupane, D.; Seok, J. Bearing Fault Detection and Diagnosis Using Case Western Reserve University Dataset with Deep Learning Approaches: A Review. *IEEE Access* **2020**, *8*, 93155–93178. [\[CrossRef\]](#)
89. Farag, M.M. Towards a Standard Benchmarking Framework for Domain Adaptation in Intelligent Fault Diagnosis. *IEEE Access* **2025**, *13*, 24426–24453. [\[CrossRef\]](#)
90. Idrees, A.K.; Alhussaini, R.; Salman, M.A. Energy-efficient two-layer data transmission reduction protocol in periodic sensor networks of IoTs. *Pers. Ubiquitous Comput.* **2023**, *27*, 139–158. [\[CrossRef\]](#)
91. Aldahhri, E.A.; Almazroi, A.A.; Alkinani, M.H.; Ayub, N.; Alghamdi, E.A.; Janbi, N.F. Smart Farming: Enhancing Urban Agriculture Through Predictive Analytics and Resource Optimization. *IEEE Access* **2025**, *13*, 72375–72388. [\[CrossRef\]](#)
92. Lynda, D.; Brahim, F.; Hamid, S.; Hamadoun, C. Towards a semantic structure for classifying IoT agriculture sensor datasets: An approach based on machine learning and web semantic technologies. *J. King Saud Univ.-Comput. Inf. Sci.* **2023**, *35*, 101700. [\[CrossRef\]](#)
93. Long, S.; Miao, F. Research on ZigBee wireless communication technology and its application. In Proceedings of the 2019 IEEE 4th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), Chengdu, China, 20–22 December 2019; pp. 1830–1834. [\[CrossRef\]](#)
94. Yang, X.; Zhang, L.; Shu, L.; Jing, X.; Zhang, Z. SILF Dataset: Fault Dataset for Solar Insecticidal Lamp Internet of Things Node. *Sensors* **2025**, *25*, 2808. [\[CrossRef\]](#) [\[PubMed\]](#)
95. Huang, K.; Shu, L.; Li, K.; Feng, Y.; Jiang, Z.; Zhu, Y. Insecticidal counting dataset based on one solar insecticidal lamp and two cameras. *Front. Plant Sci.* **2022**, *13*, 995118. [\[CrossRef\]](#)
96. Nagaraj, D.; Proust, E.; Todeschini, A.; Rulli, M.C.; D’Odorico, P. A new dataset of global irrigation areas from 2001 to 2015. *Adv. Water Resour.* **2021**, *152*, 103910. [\[CrossRef\]](#)
97. Kushal Kumar, B.N.; Balakrishna, R.; Panduranga Rao, M.; Ashok Kumar, P.S. Comprehensive Insights into Machine Learning for Intrusion Detection Systems in IoT and its Datasets. In Proceedings of the 2024 4th International Conference on Data Engineering and Communication Systems (ICDECS), Bangalore, India, 22–23 March 2024; pp. 1–5. [\[CrossRef\]](#)
98. Wang, L.; Chi, J.; Ding, Y.; Yao, H.; Guo, Q.; Yang, H. Transformer fault diagnosis method based on SMOTE and NGO-GBDT. *Sci. Rep.* **2024**, *14*, 7179. [\[CrossRef\]](#) [\[PubMed\]](#)
99. Ma, X.; Gu, M. A Bayesian FMEA-Based Method for Critical Fault Identification in Stacker-Automated Stereoscopic Warehouses. *Machines* **2025**, *13*, 242. [\[CrossRef\]](#)
100. Li, Z.; Shen, S.; Liu, Z.; Chen, Y. A Novel Multisource-Domain Adaptation Framework for Bearing Fault Diagnosis Based on Adversarial Network and Feature Enhancement. *IEEE Trans. Instrum. Meas.* **2025**, *74*, 3510712. [\[CrossRef\]](#)
101. Tang, Z.; Zhang, J.; Zhang, K. What-is and how-to for fairness in machine learning: A survey, reflection, and perspective. *ACM Comput. Surv.* **2023**, *55*, 1–37. [\[CrossRef\]](#)
102. Naranjo, J.E.; Alban, J.S.; Balseca, M.S.; Bustamante Villagómez, D.F.; Mancheno Falconi, M.G.; Garcia, M.V. Enhancing Institutional Sustainability Through Process Optimization: A Hybrid Approach Using FMEA and Machine Learning. *Sustainability* **2025**, *17*, 1357. [\[CrossRef\]](#)
103. Gandhare, S.; Narad, S.; Kumar, P.; Madankar, T. Integrating Quantitative Parameters for Automating Medical Equipment Maintenance Using Industry 4.0 and FMEA. In Proceedings of the 2024 2nd DMIHER International Conference on Artificial Intelligence in Healthcare, Education and Industry (IDICAIEI), Wardha, India, 29–30 November 2024; pp. 1–6. [\[CrossRef\]](#)
104. Wang, H.; Wang, X.; Yang, Y.; Gryllias, K.; Liu, Z. A few-shot machinery fault diagnosis framework based on self-supervised signal representation learning. *IEEE Trans. Instrum. Meas.* **2024**, *73*, 3509114. [\[CrossRef\]](#)
105. Jiang, Q.; Lin, X.; Lu, X.; Shen, Y.; Zhu, Q.; Zhang, Q. Self-supervised learning-based dual-classifier domain adaptation model for rolling bearings cross-domain fault diagnosis. *Knowl.-Based Syst.* **2024**, *284*, 111229. [\[CrossRef\]](#)
106. Hu, H.; Feng, Z.; Li, R.; Ma, Y.; Yang, S. A Novel Cross-Sensor Self-Supervised Learning Method for Rotating Machinery Fault Diagnosis. In Proceedings of the ICASSP 2024—2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Seoul, Republic of Korea, 14–19 April 2024; pp. 7795–7799. [\[CrossRef\]](#)
107. Zhao, Y.; Sun, Z.; Wang, Q.; Cui, D.; Liu, P.; Wang, Z. Toward Electric Vehicle Safe Operation: Unsupervised Learning-Enabled Multiple Fault Diagnosis in Lithium-ion Battery Systems. In Proceedings of the 2024 IEEE 10th International Power Electronics and Motion Control Conference (IPEMC2024-ECCE Asia), Chengdu, China, 17–20 May 2024; pp. 2812–2817. [\[CrossRef\]](#)

108. Ghosh, A.; Hong, J.; Yin, D.; Ramchandran, K. Robust Federated Learning in a Heterogeneous Environment. *arXiv* **2019**, arXiv:1906.06629. [[CrossRef](#)]
109. Sravani, M.; Purbey, S.; Chakradhar, B.; Kumar Choudhary, A. Design of an Efficient Multidomain Augmented Data Aggregation Model to Solve Heterogeneity Issues for IoT Deployments. In Proceedings of the 2023 2nd International Conference on Applied Artificial Intelligence and Computing (ICAAIC), Salem, India, 4–6 May 2023; pp. 1330–1336. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.