



OPEN ACCESS

EDITED BY

Kunpeng Zhu,
Chinese Academy of Sciences (CAS), China

REVIEWED BY

Xin Lin,
Wuhan University of Science and Technology,
China
Yanfeng Peng,
Hunan University of Science and Engineering,
China

*CORRESPONDENCE

Yassmin Seid Ahmed,
✉ yassmin.seidahmed@kfupm.edu.sa

RECEIVED 22 January 2025

ACCEPTED 31 March 2025

PUBLISHED 17 April 2025

CITATION

Seid Ahmed Y, Abubakar AA, Arif AFM and Al-Badour FA (2025) Advances in fault detection techniques for automated manufacturing systems in industry 4.0. *Front. Mech. Eng.* 11:1564846. doi: 10.3389/fmech.2025.1564846

COPYRIGHT

© 2025 Seid Ahmed, Abubakar, Arif and Al-Badour. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](#). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Advances in fault detection techniques for automated manufacturing systems in industry 4.0

Yassmin Seid Ahmed*, Abba A. Abubakar, Abul Fazal M. Arif and Fadi A. Al-Badour

¹Mechanical Engineering Department, King Fahd University of Petroleum & Minerals, Dhahran, Saudi Arabia, ²Interdisciplinary Research Center for Advanced Materials, King Fahd University of Petroleum & Minerals, Dhahran, Saudi Arabia

Fault detection and diagnosis are essential for maintaining the continuous operation of manufacturing systems. To achieve this, an innovative tool is required to immediately identify any faults in the production process and recommend the appropriate mechanisms to be adopted proactively to prevent future mishaps or accidents. This capability is critical for many industries to improve the efficiency and effectiveness of their production processes. Several methods can be used to detect trends or patterns in any given process and determine if the process variable is within normal limits. However, these techniques may only detect evident process characteristics or defects while leaving behind latent ones. This paper aims to review recent achievements and classics in fault diagnosis and detection, and suggest steps that can be taken to plan and implement this process. It will also explore emerging research streams, critical issues in the field, and strategies that can be applied to overcome these barriers. The paper outlines how the performance of fault detection and diagnostics can be improved in production processes and how a safer and fully efficient production environment can be promoted.

KEYWORDS

fault detection, production processes, signal acquisition, fault diagnosis, industry 4.0

1 Introduction

Over the past three centuries, industrial productivity has significantly increased. Machines and processes were innovated and improved in the first and second industrial revolutions, with the first based on mechanical technologies employing steam and water and the following on electricity and complex machine tools (Bilbao-Ubillos et al., 2024). The third industrial revolution, which started in the 1950s, was more digitized, with semiconductors and communication networks, allowing automated manufacturing (Miah et al., 2024). The production business has experienced the positive influence of artificial intelligence (AI) and machine learning (ML) technologies since their advent 10 years ago. That has influenced the growth of productivity levels, resource consumption and waste reduction, and the strengthening of sustainability, worker safety, and quality and output (Gawde et al., 2024). An AI foundation provides a promising basis for complex manufacturing processes, including fault detection and diagnosis (FDD) techniques. AI methodologies enable manufacturers to identify and resolve operational obstacles in real

time. This means the production process is less prone to bottlenecks, resulting in higher-quality products and improved company performance. Therefore, organizations should analyze how AI technology can empower production and make them one of the successful players in the market.

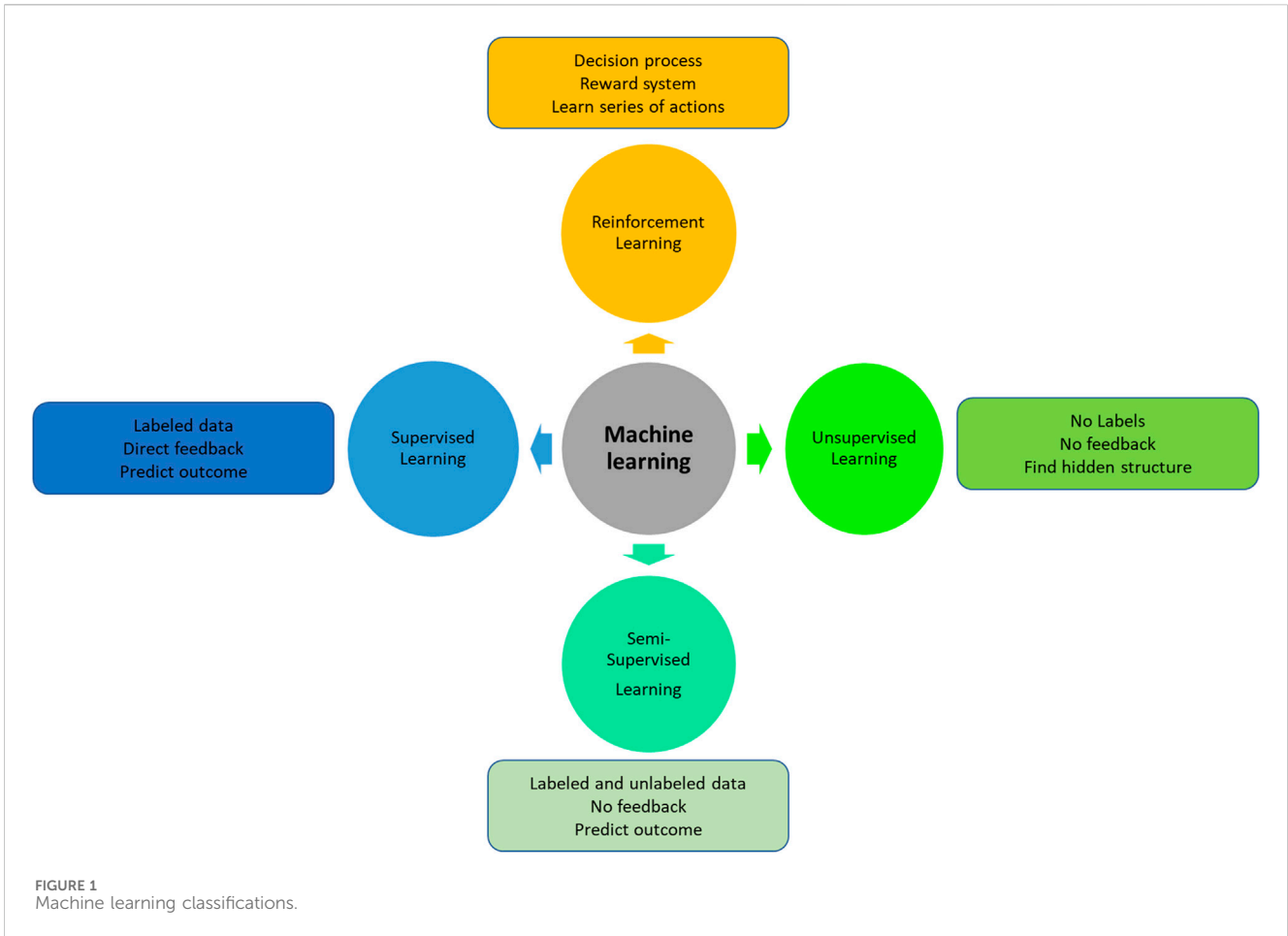
1.1 Transforming industry 4.0 with machine learning

The future “Industry 4.0” revolution aims to foster a manufacturing environment that is intelligent, self-dependent, interoperable, and operates in a real-time mode (Gawde et al., 2024). To achieve the goals of Industry 4.0, it is essential to have advanced capabilities in computing, communications, and industrial innovation (ElMaraghy and ElMaraghy, 2022; Bilbao-Ubillos et al., 2024). The Internet of Things (IoT) also allows for secure, dependable, and instantaneous interconnectivity among various devices, such as sensors, machines, robots, and actuators. The IoT comprises numerous communication networks, including machine-to-machine deployments, cloud technologies, 5G networks, and Wi-Fi networks (cloud, fog, edge, etc.) (ElMaraghy et al., 2021; Attaran et al., 2024). Despite the increased utilization of AI in decision-making, smart devices, and virtual reality (VR)/auto regressive (AR) capabilities, humans should still play a significant role in Industry 4.0 (ElMaraghy et al., 2021; ElMaraghy and ElMaraghy, 2022). Industry 4.0 promises to revolutionize current industrial production methods, benefiting companies, workers, and the environment with cutting-edge technologies. Industry 4.0 offers numerous applications for agile, expert, and automated processes that simplify complex tasks and ensure high-quality results. Moreover, the most suitable applications for widespread use can be found in production and manufacturing. Developing these applications involves identifying, predicting, and preventing errors (Gawde et al., 2024).

Machine learning algorithms play an essential role in the fast and accurate detection of faults, which implies decreasing downtime by finding the affected and faulty items or parts in real time (Zaben et al., 2024). Such algorithms will help immediately detect machines containing faulty items, components, or other equipment. Accumulating vast amounts of information can facilitate the precise prediction of machine system conditions, the remaining useful life of components, and any issues. This fact will be further utilized to build a proper maintenance program, which will help prevent failures and reduce failure-related downtimes (Nelson and Dieckert, 2024; Zaben et al., 2024). Algorithms that can understand/read human actions can eliminate errors, boost productivity, and make decisions more safely. Plenty of data is necessary to integrate ML algorithms into manufacturing decision-making. Discoveries in Big Data Analytics (BDA) revealed the crucial role of technological advances in cyber-physical systems (CPS) and IoT applications, where multiple data types were created. It can be challenging to distinguish between primary data and noise, as the acquired data often contains noise from the surrounding environment that will be processed in subsequent phases. Reliable and secure real-time data transmission and processing have become increasingly critical due to the growing interconnection of subsystems (Nelson and Dieckert, 2024).

Figure 1 provides a graphical representation of the four categories of ML (Naqvi et al., 2024) and their distinguishing features. The paragraphs below describe various categories in detail. Supervised learning entails enabling an algorithm to learn from given inputs and outputs denoted by an experienced expert in some field. Applications of this approach may be found quite widely across the field of machine learning in such tasks that may be classified under classification or regression analysis in which labeled data is available (Jiang et al., 2020). Some very popular examples of supervised machine learning models offered are support vector machines (SVM) and neural networks (NN) (Jiang et al., 2020). In the case of unsupervised learning, the learning algorithm tries to search for patterns in the given data sets without any feedback from the outside. Effective methods for performing such analyses include clustering, association rules, and self-organized maps (Botvinick et al., 2019; Alzubaidi et al., 2021). One popular approach for monitoring is principal component analysis (PCA), a technique that may be found often under the unsupervised category (Alzubaidi et al., 2021). Semi-supervised learning is a learning mechanism using labeled data to guide the training of a model. At the same time, unlabeled data are also utilized to fine-tune the model within the underlying data distribution. The key difference between supervised learning, wherein the model is allowed to be aware of the correctness of its predictions on unlabeled data while it is still being trained, and semi-supervised learning, wherein there is no feedback while using unlabeled data, is that the latter shows no such mechanism (Jiang et al., 2020). The main purpose of reinforcement learning (RL) is to distinguish between actions that yield a benefit and actions that do not according to some performance criterion already known (Wang and Biljecki, 2022). The process during which various steps are performed requires that the testing results of each possible outcome be confirmed first, after which successive steps are allowed (Botvinick et al., 2019).

Deep Learning (DL) imitates the human mind with an architecture model with several hidden layers. Such networks are considered a sturdy foundation for popular implementations of advanced ML methods. Alternatively, a perceptron is a network containing only one hidden layer (Sarker, 2021). An artificial neural network (ANN) typically has three layers: an in, hidden, and output layer. An artificial neuron is comprised of three components: the input, a weight that corresponds to each input, a bias applied to the weighted sum of inputs, and an activation function that outputs results correctly (Sedej et al., 2022). Autoencoders are ANNs that follow a particular type of architecture - encoder-decoder neural networks. They are built with the consolidation of information without human supervision, in which the model reduces the dimensionality of the data and tries to map the low-dimensional latent space. Autoencoders can shine in the domain of dimensionality reduction tasks, and thus, PCA is seen as a best practice approach to a problem that used to be a domain of methods such as PCA. Dimensionality reduction is only one of many autoencoder capabilities; the other applications are compression, denoising, feature detection, and anomaly detection. For instance, decoder reconstruction can work to assess how well an encoder performs or even generate reinforced data. Autoencoders are predominantly used in these fields as they are seen increasingly in cybersecurity, facial recognition, and image processing (Sarker, 2021). Computer vision (CV) applications that require a pattern



recognition process on image-based data tend to rely on the convolutional neural network (CNN) due to its incredible effectiveness and wide application suitability among DL architectures. Input and output nodes, convolutional layers, pooling layers, and fully connected layers are elements of CNN (Yamashita et al., 2018). Through a convolutional layer, the filter or kernel, containing trainable parameters, calculates the items of the scalar product of some segmented parts of the input image or matrix, respectively. The computation of this quantity involves the most significant processing power and memory demand that a CNN uses. The output is then taken to a rectified linear (ReLU) function, which is used as an activation before the next layer. The pooling layer reduces data by replacing some values with statistics for other data in the neighborhood. The above layer is next, and it faces only lighter workloads. In a wholly coupled layer, all neurons are necessarily linked to each other, exactly like in a standard network. An activation function is then used to create output scores in the desired shape. Disregarding the computing cost relative to CNN, they have been applied in other areas, including object recognition, data classification, and translation over the past few years (Alzubaidi et al., 2021). Recurrent neural networks (RNNs) sort out time-series data, such as text and speech readings, the best since they have the memory programmed with a unique capacity to store information about the input in general and include it. Unlike most deep learning algorithms, RNNs utilize a feedback loop, which enables this memory to be retained. However, the influence of some neurons

on a large number of other neurons through the process of remembering can cause the problems of vanishing and exploding gradients, limiting the ability of RNNs to learn new information. As a solution to this complication, the Gated Recurrent Unit (GRU) and Long Short-Term Memory (LSTM) techniques have become popular, as they incorporate gates to determine which pieces of data should be retained (Zhao et al., 2024).

As a type of linear model, Support Vector Machines (SVMs) can solve both classification and regression issues. Maximizing the distance between a line or hyperplane and the closest data points allows SVMs to approximate the optimal hyperplanes or lines for distinguishing classes. SVMs can perform regression, but their primary application is in classification. Despite producing results similar to DL methods, SVMs require less computational resources and training data than their counterparts (Sarker, 2021). Decision trees are graphical structures that use specific thresholds to branch out and evaluate predictive features recursively. Although decision trees can function independently for making predictions, using a random forest, an aggregated collection of decision trees, can enhance performance and mitigate overfitting. This ensemble learning technique involves training different trees on the subsets of data or features and later fusing the outcomes. Bagging or bootstrapping could enhance the model using multiple subsets rather than just the original training data. Despite the lower accuracy of the decision trees compared with the boosting algorithms, there is a significant variation in the graduated trees

and the dataset regarding performance. The ensemble is an idea that gives more power to “weak learners,” which are simple decision trees that attain outcomes that are more accurate than random. Through the application of simple dragons, overfitting is less likely because these models are often too basic to capture the intricacies of the dataset (Sarker, 2021). One of the popular methods, the nearest neighbor technique, works by finding the closest sample to the given feature (input) in terms of distance. This technique is applied to identify new data points closest to existing ones in terms of similarity, where similarity is generally mathematically defined by Euclidean distance. Two approaches for performing nearest neighbours use brute force and K-D and ball tree algorithms. With a simple principle at its core, the closest neighbour is adaptable to address several complex issues (Guo et al., 2003). A generative adversarial network (GAN) is an unsupervised model developed to produce entirely new data samples that look very similar to the examples found in the input data. The GAN architecture comprises two DL models - a generator that produces synthetic samples and a discriminator that identifies fraudulent samples. The generator learns from the discriminator’s feedback, while the discriminator learns from the real *versus* fake labels. This mutual learning process makes the GAN highly effective. GANs are increasingly used in real-world applications where data is limited, as they can generate data points to support critical decisions. In addition to data synthesis, GANs have potential applications in image processing, music creation, and drug discovery (Goodfellow et al., 2014).

1.2 The importance of fault detection and diagnosis

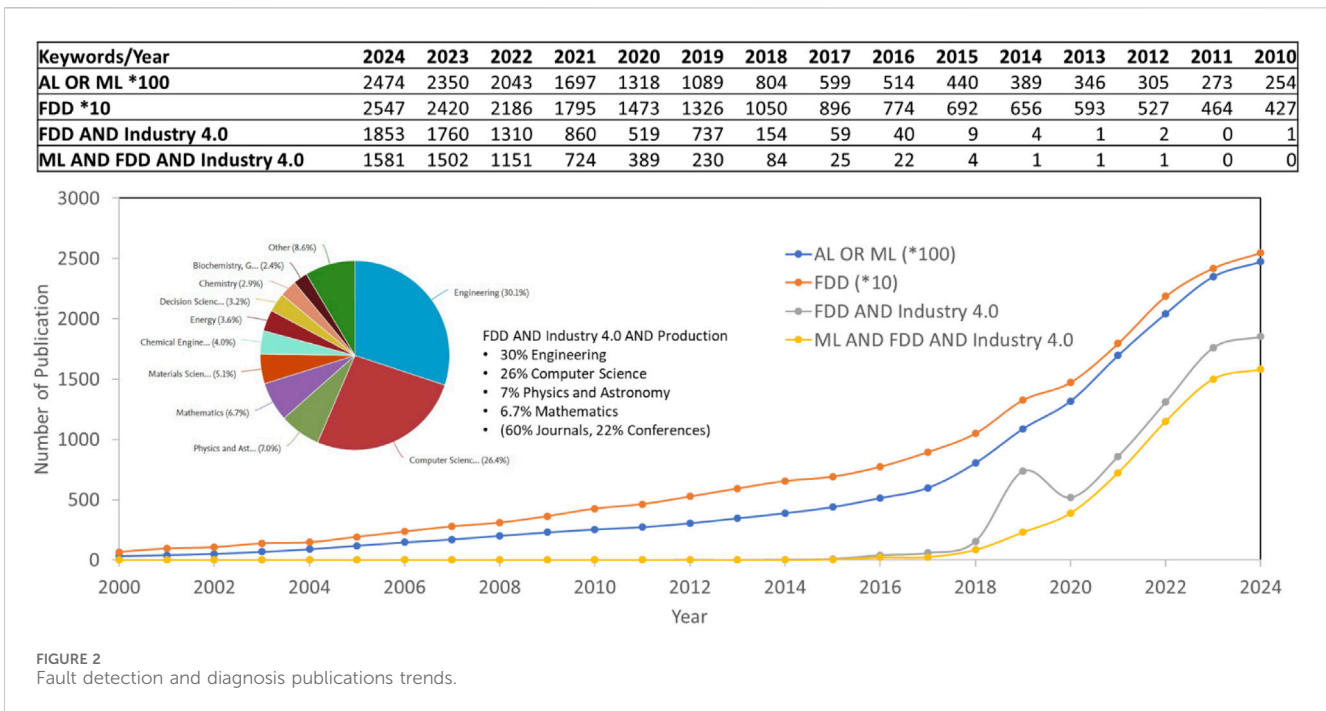
Industry 4.0 has brought about smarter manufacturing processes incorporating advanced sensors to track processes, detect faults, and prevent errors. With the increasing adoption of fully automated industrial environments, better supervision is needed to ensure optimal process efficiency, including process control and corrective actions. However, maintaining optimal performance in manufacturing processes that encounter various types of faults is a challenging task. Industries aim to enhance FDD capabilities for improved process performance (Yamashita et al., 2018; Sedej et al., 2022). For the last few decades, FDD has gained considerable attention among the industry and academia for its precious advantages like reduced cost, improved quality, and enhanced productivity of the process or product, sometimes cited in literature (Gnanamalar et al., 2023). In order to maintain a cost-effective operation, it has become more and more important to properly identify and diagnose faults since they can adversely affect product quality or throughput. In this regard, FDD is a primary yet prominent technique for process supervision based on the data gleaned from sensory tools. Different investigations proposed methods of varying kinds for either online or offline process fault detection (Mercorelli, 2024; Orhan and Celik, 2024; Shang et al., 2024). After the detection, many authors use model-based or statistical models to address the practical, theoretical, and experimental side of the problem to perform diagnosis, isolation, and identification of the root cause of process anomalies. These FDD techniques can be classified into three main groups: data-driven,

model-based, and knowledge-based methods. Typical methods involved include, but are not limited to, data-driven and model-based approaches, which have found great success in industrial process FDD due to their straightforwardness and efficiency. Real-time, online FDD is an urgent problem in some particular industries, especially for large-scale production processes.

Besides the foundational studies referred to before, several new research studies have made the field wide with new solution methods to tackle the problems of bearing failures, introducing modern different methods to deal with data imbalance, variable operating conditions, or the need for advanced AI-driven fault detection methods, such as that which (Li et al., 2024) have proposed-the adaptive clustering weighted oversampling method-to enhance the bearing fault diagnosis performance under data imbalance and variable speed operations. Due to the limited fault data available, traditional fault detection models exhibit biased learning and poor classification performance. The proposed study counteracts this limited diversity of the data by performing adaptive oversampling for the minority fault classes; further enhances the classification accuracy and robustness in a more realistic industrial environment, where it is often not easy to obtain balanced fault datasets.

Similarly, Li et al. (2023) investigate the verification of icing on wind turbine blades using a Clustered Joint Bayesian Modeling (CJBM)-based approach with imbalanced data. While mainly applied to turbine components, the manner in which the CJBM is leveraged to increase fault prediction on highly imbalanced datasets is relevant to bearing fault diagnosis. This study exemplifies how Bayesian-based models could be useful in addressing the uncertainty surrounding the requisite requirements in fault detection; it posits a framework expandable into the health monitoring of bearings facing variable environmental conditions. Moreover, Fan et al. (2024) describe the Variable Scale Multilayer Perceptron (VS-SMLP) that helps diagnose abnormal helicopter transmission system vibrations. The study shows how deep learning-based techniques applied to bearings can improve condition monitoring in rotating machinery. The model can improve feature extraction, even from noisy vibration signals, using a dynamically adjustable network scale based on the complexity of the data being processed. This offered an opportunity for more adaptive AI-driven fault detection in on-the-field applications, especially where machinery works on time-varying conditions. Furthermore, Li et al. (2024) present a semisupervised fault diagnosis approach utilizing CJBM for imbalanced bearing failure data. This study addresses the challenge of the small volume of labeled fault data in some production environments by applying semi-supervised learning, wherein both labeled and unlabeled data are used for training. The contributions of this study demonstrate that hybrid learning techniques could effectively enhance detection for early-stage faults for applications where manually labeled fault data is scarce or infeasible due to high cost. Such contributions highlight the diverse ways bearing failures could occur, thus requiring a variety of diagnosis techniques. These studies explore modern industrial emergencies regarding fault detection capabilities by exploiting such techniques as oversampling, Bayesian modeling, adaptive deep learning architectures, and semi-supervised learning.

The graph in Figure 2 displays the results of a Scopus search conducted using selected keywords in documents published



annually since 2010. According to Scopus, the number of publications related to FDD in Industry 4.0 has been steadily increasing, with the emergence of works on FDD for production in recent years. By searching for the keywords “fault detection and diagnosis AND Industry 4.0” and “fault detection and diagnosis AND Industry 4.0 AND Production” in the title, abstract, and keywords fields, over 5,635 and 3,272 results were obtained, respectively, as of November 2024. This indicates that interest in FDD in Industry 4.0 is expanding, with the involvement of researchers from different disciplines, such as engineering, computer science, and statistics. The ongoing increase in publications about FDD concerning the production process means that the field has become the subject of constant upgrades and innovations by researchers to improve the precision and efficiency of FDD. This trend will likely be the future as the demand for dependable and efficient systems is very high in various industries.

1.3 Motivation, novelty and contribution

As we continue to adopt Industry 4.0 technologies, we realize the importance of having automated FDD systems. This review paper examines the current focus on FDD in production systems. By utilizing robotics automation for work operations, production units can save on costs and increase competitiveness. We present a systematic approach to addressing errors that arise in the Industry 4.0 environment.

Compared to these diagnostic studies, such as (Mercorelli, 2024; Orhan and Celik, 2024; Shang et al., 2024), this review is intended to provide a wider and integrative basis for diagnosing and detecting faults in production environments. Most of the previous works were targeted toward certain types or diagnostics sectors, limiting them to

the particular needs of production processes. Mercorelli advanced many intelligent algorithms for fault detection in 2024; however, he focused mainly on developing algorithms rather than on their validation through extensive real-life industrial applications. It poses difficulties when one attempts to assess how feasible these schemes would be in actual implementation within real-world industrial setups, where computational limitations and adaptability matter. In turn, Orhan and Celik offer helpful insights into the marine machinery systems in 2024; however, their study is still limited to a single industry and does not conduct any comparative analysis with other FDD approaches. This industry-based approach restricts its applicability in broader production systems with diverse failure modes and operational conditions. Shang et al. lent guidelines industry-based for battery systems in 2024, but they did not deal with how FDD techniques would be able to cross multiple domains. Their work does not consider hybrid approaches incorporating data-driven, model-based, and knowledge-based approaches, a necessity increasingly felt in contemporary industrial systems.

The present review aims to bridge these gaps while offering a comprehensive account of FDD techniques’ applicability for several industrial sectors. Unlike previous studies that looked into isolated methodologies or sector-specific applications, this study comparatively evaluates data-driven, knowledge-based, and model-based approaches with a systematic assessment of their underlying efficacy in various operational environments. Furthermore, case studies from the real world are also integrated into this analysis to validate the applicability of various FDD methods practically. This will keep our conclusions benefiting theoretical and practical relevance for researchers and practitioners who develop adaptive FDD in Industry 4.0 scenarios.

The main contributions of the current work are as follows:

- It provides an in-depth analysis of three FDD strategies: data-driven, model-based, and knowledge-based. These strategies address the issues in FDD in several ways.
- It contrasts various FDD approaches, comparing their performance level, degree of complexity, and computational needs. It is meant to guide researchers and practitioners experimenting with feature detection in their specific cases.
- It presents evaluation metrics to confirm the process of FDD techniques for researchers and practitioners who want to compare different FDD methods.
- It addresses the utilization of FDD procedures in several industries. It argues that FDD is the best way to significantly increase the industry's security, dependability, and productivity.
- Unlike previous studies focusing on single methodologies or industry-specific applications, our review integrates a cross-industry perspective, allowing for broader applicability of FDD techniques in diverse production environments.
- “By incorporating empirical case studies and performance evaluations, this work ensures the practical relevance of FDD methods, addressing the real-world deployment challenges overlooked in prior research.

This work contributes to the field by providing a broader framework encompassing data-driven, model-driven, and knowledge-driven FDD methods. While previous studies have focused mainly on individual approaches, our work systematically compares these techniques, highlighting their relative advantages and limitations. To substantiate the effectiveness of these methods, empirical studies and real-world case studies have been integrated throughout the manuscript. In particular, [Section 3.4](#) presents a comparative analysis of different signal analysis techniques applied to fault detection, demonstrating their performance in various industrial scenarios. Additionally, [Section 5](#) provides empirical evidence of how data-driven, model-based, and knowledge-based methods have been implemented in practical applications, showcasing key performance indicators such as detection accuracy, false alarm rates, and computational efficiency. These contributions ensure a comprehensive evaluation of FDD methods, reinforcing their practical relevance in industrial applications.

1.4 Paper structure

This research review provides valuable insights into FDD, covering the latest automation and Industry 4.0 techniques. It is an essential introduction to FDD systems, including the industry's most used methods and elements. [Figure 3](#) illustrates the typical steps involved in FDD development, which this paper discusses as follows. [Section 2](#) explores the development of FDD, while [Section 3](#) discusses signal acquisition and analysis procedures. [Section 4](#) discusses some industrial applications of FDD. [Section 5](#) categorizes faults and discusses techniques for identifying and resolving them. [Section 6](#) highlights some directions for future work, and finally, chapter eight concludes the paper.

2 Fault detection and diagnosis development

2.1 System structure and redundancy methods

The methods for detecting and diagnosing faults in a system depend on how the system is modelled and represented; for instance, an empirical model (Patel, 2020) or a mathematical function (Bighamian et al., 2015) can clearly understand the system's dynamics. In other cases, visual representations of system activity (Collins and Woodruff, 2018) or implicit knowledge representations, such as a knowledgebase, ANN, or expert system, can be used to guide output decisions based on rules and heuristics (Zhang and Zhao, 2017; Mostafa et al., 2018; Wu and Zhao, 2020). [Figure 4](#) illustrates various system structure techniques that can be utilized in FDD development.

Redundancy is an essential aspect of the functionality of FDD systems as it plays a crucial role in enhancing their reliability and availability. Redundancy types like physical redundancy, analytical redundancy, software redundancy, and structural redundancy have different approaches to achieving this. All types involve duplicating several components from the system and incorporating backup components to support these original components in case of failure. Planning and implementing redundancy strategies properly ensures that FDD systems have dependability and resilience (Kizza, 2020). Physical redundancy consists of duplicating physical system components, whether in the form of server, actuator or sensor duplication, which can operate in parallel to provide a backup capability at the hardware level (Ahmed and Amorim, 2025). Analytical redundancy uses mathematical modeling and algorithms to detect errors by analyzing and comparing expected output and actual output. It is mainly used in control systems and mercennial behaving of processes for information checking for consistency between real and expected behaviors of the system. Software redundancy duplicates software components or processes, usually through hot standby or load-balancing techniques. Structural redundancy involves the adding of components to a system to increase its strength or capacity.

This is common in engineering structures, but can also apply to software systems with multiple modules or subsystems, to ensure that not all the modules fail for the same set of input data (Kizza, 2020). By increasing the complexity and diversity of the system, structural redundancy can protect against catastrophic failure caused by a homogenous set of incoming data. The effectiveness of each type of redundancy depends on the specific system and the types of failures it is designed to tolerate. Physical redundancy is often more expensive as compared to software redundancy. On the other hand, implementing and managing analytical redundancy can be complex due to the need for sophisticated algorithms and models. An effective system must exhibit trade-offs between redundancy, performance, and cost.

2.2 Feature extraction

Feature extraction is a crucial step in the FDD process. It involves transforming raw equipment/process data into a set of

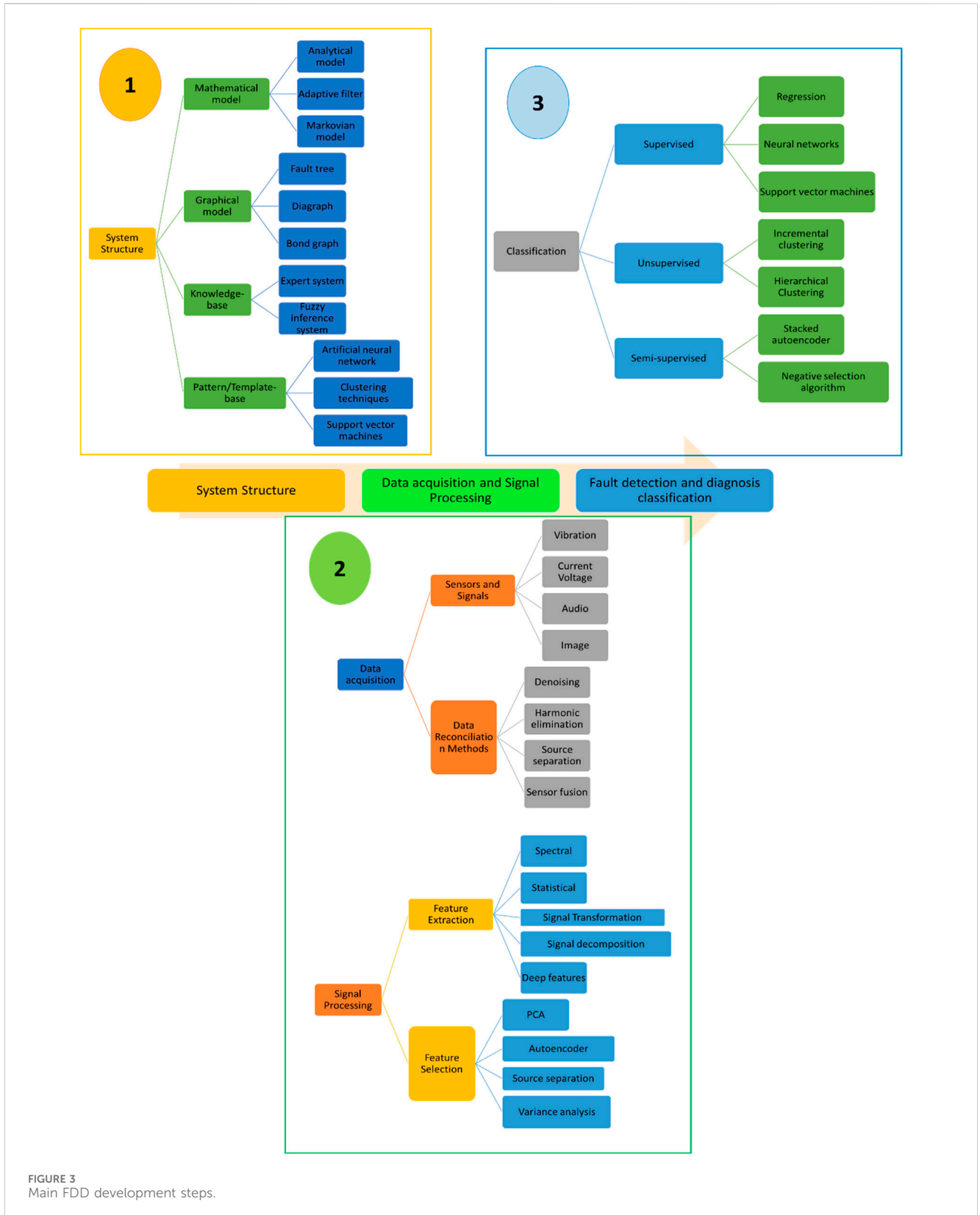
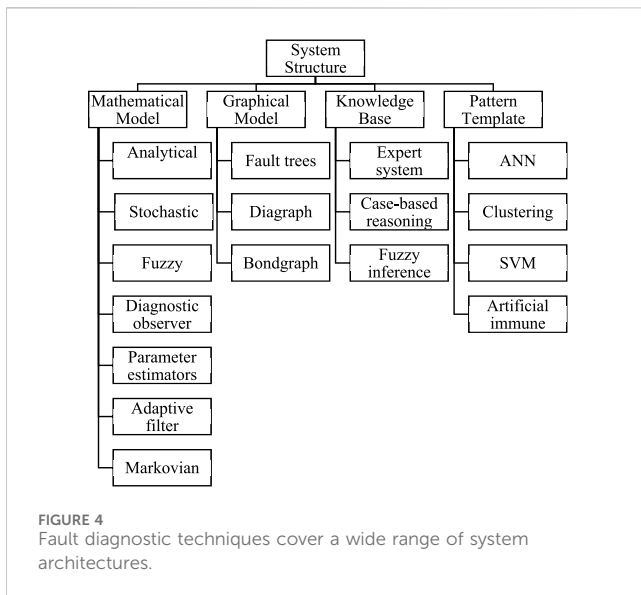


FIGURE 3 Main FDD development steps.

meaningful features that can be used to differentiate between normal and faulty system conditions. Fault detection in a system involves analyzing various features extracted from the system signals, which convey valuable information regarding the system condition. In

different fault detection and diagnosis studies, features such as mean, standard deviation, variance, and so on were extracted by denoting time domain signal processing. The kurtosis provided by (Wang and Hu, 2024) affords applicability in bearing fault detection



and insensitivity to load variations. Other highlights which show potential for help in indicating fault occurrence in a system include entropy, homogeneity (recommended by (Benmoussa and Djeziri, 2017), energy kurtosis (proposed by (Wang and Lee, 2013), and power efficiency (Chen and Kamara, 2011). Several models and approaches are brought into play for estimating attributes that characterize faults as discussed throughout the literature, such as autoregressive system modeling (Yan et al., 2014), nonlinear regression (Zheng et al., 2016), and parameter estimation (Bouchikhi et al., 2015). Some proposed using Park's transformation for feature extraction in machine diagnosis (Kwak et al., 2015). The broad, dominant harmonic amplitude and phase may be computed analytically without employing auxiliary tools and inferences from spectral characteristics (Shah and Patel, 2014; Boudinar et al., 2015). The spectral means, crest factor, and spectral entropy reveal worth and meaning in the frequency domain statistically (Yang et al., 2024).

Chen, Rhee and Liu, 2019 stated that using vibration data, envelope analysis and parametric spectral estimation have traditionally been applied to detect defects in bearings and rotors. Spectral domains provide helpful information for specific causes of defects, but tend to be ineffective in identifying minor defects. Abid et al. (2020) developed a combined time-frequency feature extraction strategy for analyzing non-stationary events to overcome shortcomings of the traditional methods. The most common time-frequency analysis techniques include Short-Time Fourier Transform, Empirical Mode Decomposition (Camarena-Martinez et al., 2015) and Wavelet Transform (Abad et al., 2016). These techniques have provided appropriate ways of extracting a complementary set of perspectives from the data to characterise variation under different operational scenarios (Han et al., 2018; Abid et al., 2020). Further, feature dimensionality reduction and selective techniques included principal component analysis (PCA) and independent component analysis (ICA) (Huang et al., 2015). On the other hand, other methods, namely, variance analysis, feature pruning, and deep feature selection (Abid et al., 2020), can further improve fault detection.

2.3 Fault type analysis

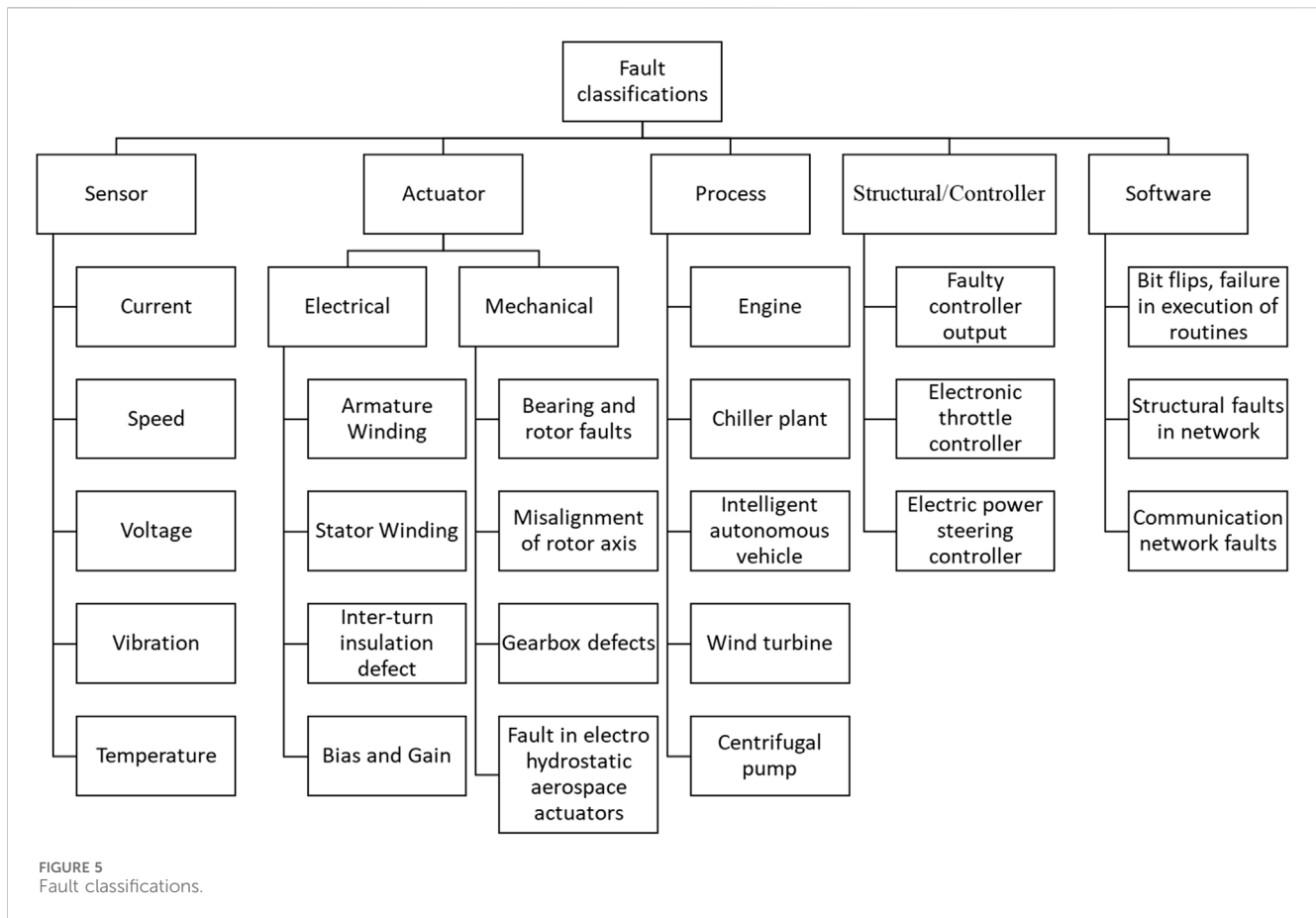
After describing the system model and fault redundancy, the type of fault present within a system is among the major considerations in selecting the appropriate type of FDD system. Figure 5 provides a summary of various faults afflicting an industrial system. It is crucial to have some knowledge of the various fault types to properly select the FDD system to be well-equipped to combat possible problems. Faults are classified based on their nature and dynamics. A transient fault appears due to a sudden change or transient effect within the system for a short period and goes away (Wu et al., 2023). However, damage caused by permanent faults cannot be reset unless the faulty component has been repaired or replaced. Intermittent faults are only active for short periods (Abid et al., 2015). These four generalized faults are hardware faults, software faults, network faults, and communication failures. They can all compromise the working of a system and lead to either performance degradation or system failure. According to many different systems, these four combined can build a really tough problem for FDD in the real world. A good FDD system should consider multiple fault types and mixed methods to detect and diagnose the problem efficiently. Hardware faults include sensor, actuator, process and structural faults (Kumar et al., 2022).

Some hardware failures involve physical component failures and more often lead to system crashes, data loss, or poorly working over time. As a result, it usually needs physical inspection, diagnostic tools, or replacement of the faulty components. Software fault comes from software, code, or logic errors and have effects like wrong outputs, infection of systems by instability, and security vulnerabilities. Corrective actions involving debugging, code analysis, and testing may also remedy possible software faults such as bit flips, runtime errors, bugs, or design flaws (Abaei et al., 2013). Network faults arise mainly due to issues of network connectivity or performance; for example, the use of incompatible protocols, unacknowledged packets, congestion, or irretrievable data loss (Kizza, 2020). Network faults can interfere with communication from one system component to another with a resultant reduction of performance and possibly system failure during engineering operations. To identify network faults, there is a need for monitoring utilities, protocol analyzers, and metrics for network performance. Communication faults occur due to the incorrect transmission of messages or interpretation between system components; examples include protocol errors and data corruption (Kizza, 2020). Communication faults may be examined by analyzing the communication protocol, error detection mechanisms, and data integrity.

3 Signal acquisition and analysis

3.1 Data acquisition

The decision to use an FDD scheme is influenced by the amount of readily accessible system/process data. FDD systems rely heavily on data acquisition (DAQ), which mainly involves capturing real-world signals, converting them into digital format, and storing them for analysis (Cho et al., 2018). DAQ tools are crucial for gathering the raw data necessary for fault detection and diagnosis systems.



DAQ systems consist of various hardware and software tools, including sensors, signal conditioning modules, analog-to-digital converters (ADCs), software, databases, programmable logic or microcontrollers (PLCs), network interfaces and DAQ boards (Cho et al., 2018). Sensors convert physical parameters (temperature, pressure, vibration, etc.) into electrical signals. DAQ boards are specialized hardware devices that connect to a computer and provide various input/output channels for analog and digital signals. Signal conditioners or sensor modules are devices that prepare the raw sensor signals for acquisition by amplifying, filtering, or converting them to suitable levels compatible with DAQ systems. DAQ tool libraries developed with general-purpose programming languages, such as PyDAQmx (Python), DAQmx C API (C/C++), Raspberry/Arduino Uno, NI LabVIEW, and MATLAB/Simulink, can all be used to collect custom data for FDD systems. For storing and managing large volumes of acquired data, databases like MySQL, PostgreSQL, or SQL Server are often used. When selecting DAQ tools for an FDD system, it is essential to consider factors such as sensor types and compatibility, sampling rate and resolution, data storage and transfer media, synchronization of multiple data sources, and data preprocessing. Even though extensive data is needed for accurate FDD analysis, online system health monitoring can be effectively carried out using a small amount of data by incorporating *a priori* knowledge, such as a threshold value (Salehifar et al., 2014; Sadeghkhanian et al., 2018; Su and Chen, 2019), or fault decision indicators (Bouchikhi et al., 2015). The threshold or fault indicators can pinpoint anomalies in any

system or engineering operation. As a result, the availability of vast amounts of historical process data and the widespread adoption of sensor networks for monitoring and assessing process variables have resulted in increased transparency and efficiency in industrial settings (Cheng et al., 2019).

3.2 Signals analysis

Multiple sensors detect a wide range of physical quantities in an FDD system. These sensors provide essential information about the system's normal and faulty states. Signal processing and analysis tools are necessary for extracting meaningful information from raw sensor data in fault detection and diagnosis systems. These tools help to identify anomalies, trends, and patterns that indicate potential equipment failures. Different methods of monitoring machine tools/equipment for malfunction detection exist. Figure 6 shows the signals commonly used for FDD, further discussed below. The analysis of vibration and current signals are two of the most prevalent approaches. However, acoustic emission and image data also provide useful information for FDD analysis.

3.2.1 Vibration

Whenever a fault happens in a machine, the dynamic behaviour of the machine will change and vibrational signals directly capture this. Cheng et al. (2019b) utilized vibration signal analysis to develop

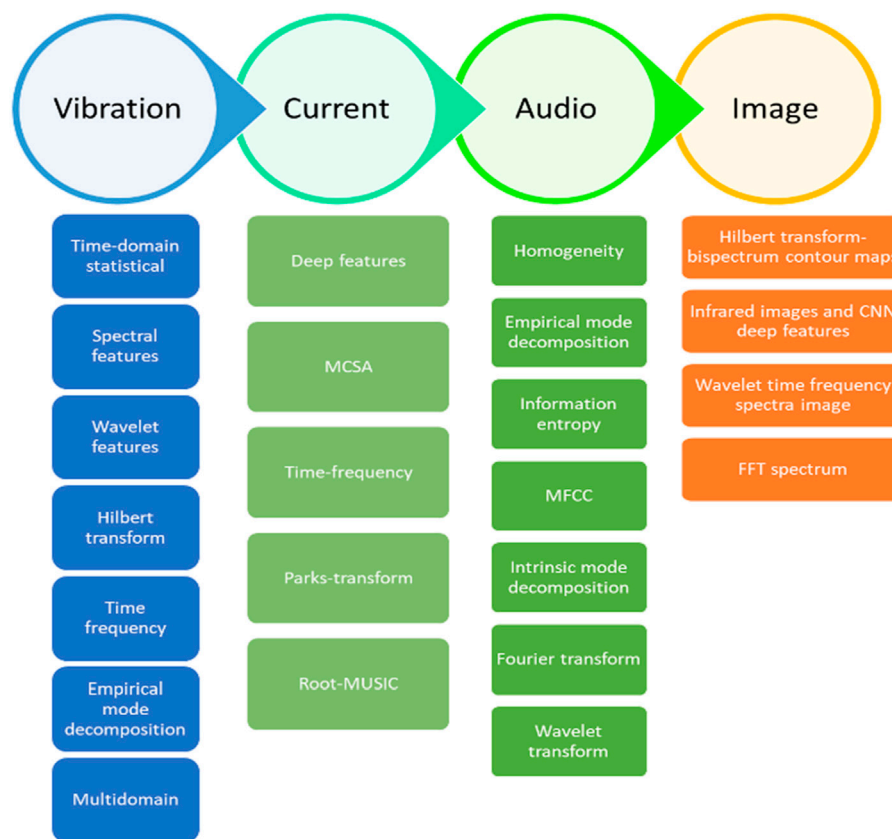


FIGURE 6
Fault diagnosis signals and features.

gearbox fault diagnosis methods. Several fault-characterizing models based on vibration signals, including those proposed in (Wang and Hu, 2024) were designed for diagnosing different types of bearing faults. As shown in Figure 6, time-based, frequency-based, and time-frequency-based features can be extracted from vibration signals, which are usually captured using vibration sensors (e.g., accelerometers) and specialized signal processing software (e.g., MATLAB/Simulink, Simcenter Testlab, Brüel and Kjær PULSE, etc.), making vibration signal analysis a mature area for fault diagnosis. Common signal processing algorithms/techniques used for vibration data filtering and analysis include Fast Fourier transform (FFT), wavelet transform, Hilbert transform, envelope analysis, statistical analysis (e.g., kurtosis, crest factor, etc.), and machine learning (e.g., SVM, neural networks, etc.). Although vibration signal-based fault diagnosis methods exhibit high accuracy in most cases, they require empirical knowledge to select fault features. Recent advancements in DL led to the development of fault diagnosis methods that automatically choose features from a set of pre-established features (Xia et al., 2018; Abid et al., 2020). Many researchers devised techniques for bearing fault diagnosis, and Case Western Reserve University (CWRU) established a vibration signal dataset library for bearing defects. Fault classification accuracy for the CWRU dataset was excellent in comparative studies on bearing fault diagnosis (Boudiaf et al., 2016; Abid et al., 2020).

3.2.2 Current

Recent works show that faulty drives (Salehifar et al., 2014), winding flaws (Abbasi et al., 2018), contactor defects (Lin and Du, 2018), insulation damage, and transients (Sadeghkhani et al., 2018) can all be identified swiftly and precisely by analyzing current signals. Classifying current-based fault detection strategies as residual- or feature-based is possible. Residual-based fault detection methods rely on system models such as analytical models (Benmoussa and Djeziri, 2017), analytical redundancy relationships, Kalman filters, regression models (Yan et al., 2014), system identification models (Salehifar et al., 2014), and diagnostic observers. Other research focuses on the modeling and widespread application of fault-characterizing attributes in machine fault diagnosis (Ardali et al., 2024). Signal feature extraction methods are also used to create fault diagnosis schemes, including time-domain features (Abid and Khan, 2017), frequency-domain features, and time-frequency features (Seid Ahmed et al., 2020). Although motor current signature analysis (MCSA) (Haddad and Strangas, 2016) is the most used method due to its ease of use, it requires precise knowledge of motor slip and a suitable frequency resolution to detect and diagnose faults accurately. As a result, other techniques, including Parks transform, Root-MUSIC and DL, have recently been utilized to analyze current signals containing FDD fingerprints (Figure 6). The common software tools used for FDD analysis with current signals include MATLAB, LabVIEW, Python libraries and other libraries for power system analysis tools.

Furthermore (Salehifar et al., 2014), propose a flexible and load-independent current-signal-based fault diagnosis scheme for five-phase motor drives. Some research. Cheng et al. (2018) use current and voltage measurements for fault detection.

3.2.3 Audio

Industrial FDD techniques primarily focus on vibration and current signal analysis. Nonetheless, research on using auditory signals and acoustic machine fingerprints remains ongoing (Rubhini and Ranjan, 2017). Audio fault diagnosis techniques face a significant obstacle in noisy industrial environments due to sound signal contamination, which results in a low signal-to-noise ratio (SNR) (Kemalkar and Bairagi, 2017). To address this, denoising preprocessing methods have been developed and are practical for use in the field. Recent advances in separating and cleaning up audio have been presented by Yoshida et al. (2023). Obtaining a signal with the maximum SNR is crucial to achieve effective feature extraction (e.g., Mel-Frequency Cepstral Coefficients (MFCCs)). Unlike other sensors, audio-based techniques do not require physical modification to the machine and can be used when other sensors are not applicable (Yoshida et al., 2023). Moreover, algorithms for spectral analysis, time-frequency analysis and data classification/clustering have recently been used to identify faults in machines/equipment using acoustic data. Software tools used for audio signal processing include: Audacity, MATLAB, and Python libraries (e.g., LibROSA, TensorFlow, PyTorch, and Scikit-learn).

3.2.4 Image processing

The human sense of sight is crucial for identifying objects, with an image being the most fundamental visual representation. Advances in convolutional (CV) neural networks have led to its widespread use in image processing (Zhang, 2013; Yaman et al., 2015). A system's health status can be detailed by processing an image with metadata. Image processing techniques have been employed for detecting and characterizing structural abnormalities (Yaman et al., 2015). Common techniques involving image segmentation, feature extraction, object detection and recognition and CV neural networks are heavily used to detect anomalies in equipment/process data. Yaman et al. (2015) propose an edge detection and segmentation technique using image processing to identify flaws in rail surfaces. Infrared images are used for fault diagnosis based on temperature changes (Liu et al., 2017). Image processing is also being explored as an alternative diagnostic method for machine vibration signals. The two-dimensional computing approach provides automatic feature extraction, which is superior to manual feature selection but limited by the one-dimensional signal analysis approaches (Abad et al., 2016). Present an image processing method for feature extraction from wavelet time-frequency spectra images obtained from motor vibration signals. An alternative method for fault diagnosis is proposed by (Wang and Cheng, 2016), who suggest demodulating the vibration signal with a Hilbert transform and then using the resulting bispectrum image to describe its properties. Experiments showed that image processing techniques are effective for fault analysis based on vibration signals.

3.3 Reconciliation techniques

Automatic adjustments to measurements can be made using system data and mathematical models with the help of methods for data validation and reconciliation (DVR). There will always be some degree of error in the measured system data because measurements are obtained using imperfect instruments with different degrees of accuracy. Furthermore, errors can occur during data transmission due to runtime issues like mapping failures, transformation logic failures, and more. Inaccurate readings can also result from miscalibration, faulty instrument installation, and power fluctuations. Failure to reconcile faulty measurements can lead to data outages, incorrect values, duplicated records, or erroneous information, severely compromising data quality and leading to incorrect insights and diagnoses that go unnoticed by system users. The following is a brief discussion of the various data reconciliation methods described in the literature in the context of FDD. Wavelet transform-based (WT) vibration signal denoising techniques were presented to locate tooth faults in gearboxes and alternator faults (Abad et al., 2016). With the help of empirical mode decomposition and Fast Fourier Transform (FFT), Chen et al. (2019) applied a denoising technique to the vibration signal. For wind turbine fault detection, the autoencoder denoising approach was implemented; this method cancels out measurement noise and accounts for unexpected disturbances (Kumar et al., 2022). Harmonic error in position estimation is minimized in synchronous motor drives by employing a quadrature phase-locked loop (PLL) tracking reckoner in conjunction with an adaptive filter based on a recursive least squares method (Wang et al., 2014). Researchers traditionally proposed model-based methods for dealing with the difficulty of multiple source separation in audio signal processing (Wang et al. (2012) investigated maximum likelihood estimation for missing data in mixed signals as another way of source separation. In (Huang et al., 2015), a blind source isolation method known as Fast ICA is applied to vibration signals generated by a faulty ball screw.

On the other hand, various industrial systems, such as fish processing machines (Han et al., 2018), mobile robots (Abid et al., 2015) and transportation industry equipment (Nguyen et al., 2018), are now using multiple sensor fusion methods to enhance data quality. Information fusion techniques employing the Dempster-Shafer evidence theory (Jiang and Cao, 2016) and weighted average of sensor confidence levels (Abid et al., 2015) were also introduced. Majumder and Pratihar (2018) have presented a fuzzy sensor fusion approach to tackle sensor drift and intermittent failures. However, determining which sensors are accurate and which are not poses an extra challenge for data reconciliation. In the literature, several methods, such as the sensor confidence value and the fuzzy membership function, were suggested for identifying and avoiding incorrect sensor data. In the context of fault diagnosis, a method employing multiple separate extreme learning machines (ELMs) with a majority (Jiang and Cao, 2016) voting system was developed for oil-filled transformer fault diagnosis (Zhang and Zhai, 2019) Another hybrid strategy for bearing fault diagnosis was proposed by (Zhou et al. (2018) which used a cloud similarity measurement algorithm with weighted voting.

3.4 Comparison of signal analysis methods for specific failures

This section highlights a detailed comparison of the various signal analysis methods used for FDD concerning detection performance on specific failure types, such as faults in bearings, gears, and motors. Through qualitative insights and quantitative performance parameters, such as accuracy of identification of faults, false alarm rates, and computational needs, guidance in selecting a method most suitable for a given application will be provided. However, detailed quantitative results and case studies supporting these comparisons will be presented in Chapter 4.

3.4.1 Vibration-based methods

Among other options, many researchers refer to vibration analysis for FDD, particularly for bearing faults. Vibration signals are captured by accelerometers and processed according to techniques such as FFT, wavelet transform, or Hilbert transform. It has been reported that these vibration-based methods have the potential to detect with over 90% accuracy under controlled conditions (Pan et al., 2019). For instance, Tabrizi et al. (2015) used wavelet packet decomposition to enhance the detection of early-stage bearing defects by isolating fault-specific frequency bands. Yet, this approach suffers from manual selection of features based on empirical knowledge and also passes through the extremely noisy environment, which hurts the overall performance. Advancements in recent advanced techniques that involve deep learning (Mumuni and Mumuni, 2024) are beginning to automate the extraction of various features, thus mitigating the dependence on manual identification of tuning parameters.

3.4.2 Current-based methods

Current signal diagnosis is one of the most efficient ways to diagnose motor faults. Anomalies associated with motor fault types are detected via analysis of electrical current waveforms, using methodologies such as the Motor Current Signature Analysis (MCSA) and residual-based methods. It has been emphasized in the research by (Milles et al., 2024) that current-based methods could, in fact, more rapidly identify burning-out windings, contactor defects, and many other failure types. Furthermore, Guan et al. (2024) reported that the new ways of extracting features from the current signals result in a better rate of fault detection for motors. Nevertheless, the performance of these techniques relies heavily on precision calibration, i.e., determination of motor slip and frequency resolution. It may be susceptible to increasing false alarm mechanisms in varying loads.

3.4.3 Audio-based techniques

When conventional sensor modalities become challenging to implement, audio-based FDD methods provide an alternative. They stake the machinery's acoustic emissions and analyze the signals with algorithms capable of extracting some features, including MFCCs. Hu et al. (2025) showed that with sophisticated denoising algorithms, it is possible to improve the SNR in a noisy industrial environment dramatically. Yet, audio-based techniques tend to show less detection accuracy because isolating fault-specific acoustic signatures from the ambient background

noise is intrinsically more challenging, as (Harandi et al., 2025) pointed out.

3.4.4 Image processing methods

Since image processing techniques engaged with convolutional neural networks (CNNs) have effectively detected gear and structural faults, fantastic information in the form of visual inputs in regular cameras or infrared always aids in depicting the details of features to explain anomalies. For example, Haroon et al. (2024) segmented the rail surfaces and used edge detection to detect surface defects. (Liu et al. (2012) showed that the time-frequency spectra images synthesized from motor vibration signals could be analyzed for fault detection. Though these techniques provide high interpretability and diagnosis, they usually demand significant computational resources and hefty preprocessing, hampering their utilization in real-time applications.

3.4.5 Quantitative comparative analysis

The comparative performance summary of these approaches can be seen in Table 1. The table includes primary metrics that express performance, such as detection accuracy, false alarm rates, calculation complexity, and more practical considerations such as data quality and noise robustness. This evaluation, while affirming that vibration-based methods perform best for the detection of bearing faults, indicates that for motor faults, current analysis is ideal; in addition, image processing techniques provide the best diagnostic detail for gear faults. Evaluation of the trade-offs between computational cost and detection sensitivity is done critically, emphasizing the importance of deciding the most suitable technique depending on the type of failure and operating conditions.

The comparison study illustrated in detail in Chapter 4 shows that all applied signal analysis tools exhibit specific benefits and shortcomings. There is no method which can be considered acceptable in general; instead, by the characteristics of the fault of interest, e.g., bearing, gear, or motor, and by taking into consideration practical aspects like data quality, computational cost, and other environmental factors, the correct choice will be made by the user. Therefore, such an insight into the actual performance provides a real opportunity for design among practitioners of sufficient robust FDD systems, which are not around the corner but are generally required by modern industry.

4 Applications of fault detection and diagnosis in industry

Fault detection and diagnosis (FDD) has a broad array of benefits in industries for sustaining operational efficiency by identifying and understanding system and equipment failures (Khireddine et al., 2014; Lin and Ghoneim, 2016). Advanced FDD systems of today, with the use of technologies such as machine learning, data analytics, and sensor networks, monitor performance on a real-time basis to identify deviations in performance, pinpoint root causes, and predict impending failures. The proactive approach views this action as one that extends the life of the equipment, optimizes the production process, reduces operation costs, and maintains safety levels. That

TABLE 1 Comparative analysis of signal analysis methods for fault detection.

Method	Typical failure types	Advantages	Disadvantages
Vibration-Based	Bearing, gear faults	• High detection accuracy (often >90%)	• Requires empirical feature selection
		• Mature techniques using FFT, wavelet, and Hilbert transforms	• Sensitive to ambient noise
		• Effective feature extraction across time, frequency, and time–frequency domains	• May need manual tuning
Current-Based	Motor faults	• Rapid fault detection	• Highly dependent on precise parameter calibration (e.g., motor slip)
		• Effective for diagnosing electrical anomalies in motors	• May exhibit higher false alarm rates under variable loads
		• Supports residual-based and feature-based approaches	
Audio-Based	General machine faults: when physical sensors are impractical	• Non-invasive and cost-effective	• Susceptible to environmental noise
		• Advanced denoising algorithms can improve SNR	• Generally lower detection accuracy compared to vibration/current analysis
		• Useful in scenarios where traditional sensors are limited	
Image Processing	Gear faults, structural anomalies	• High interpretability and detailed fault characterization	• High computational cost
		• Automatic feature extraction via CNNs	• Requires extensive preprocessing
		• Suitable for visual inspection of defects	• May not be suitable for real-time applications

is the philosophy of Smart Manufacturing and Industry 4.0. (Bilbao-Ubillos et al., 2024; Miah et al., 2024).

4.1 Rotating machinery

Implementing modern FDD methods in industry has been widely acknowledged as essential for monitoring industrial processes and enhancing key performance indicators (Kankar et al., 2011; Kusiak and Verma, 2012; Liu, 2012). Specifically, the rotating machinery industry requires process and quality control tools and increased automation to enhance the throughput rate. Mechanical FDD is important because detecting early-stage faults would support maintenance scheduling, timely spare parts requisition, and avoid catastrophic failures. Rolling element bearings in rotating equipment are a leading cause of machine failure. They are challenging to detect at early stages (I and II) based on conventional vibration measurement and signal analysis methods.

Recent studies have demonstrated that machine learning-based FDD techniques significantly improve early-stage fault detection. To detect faults in bearings, Kankar et al. (2011) tested the entire panoply of traditional vibration analysis *versus* ANN and SVM models, achieving classification accuracies which surpassed 95% in favor of the latter. Also, in real-world cases, Liu et al. (2022) performed a study on industrial pump monitoring, producing about a 35% reduction in false alarms and a 28% increase in efficiency of maintenance scheduling when using SVM-based methods. Kankar et al. (2011) state that their suggested method of using artificial neural networks and SVM performed better than the traditional vibration analysis in the process of bearing fault

detection, achieving an accuracy of not less than 90%. This, in some sense, shows the usefulness of AI-based FDD in decreasing false positives and elevating fault detection in earlier stages. Kusiak and Verma (2012) conducted research using data mining to discover the root causes for failures in bearings, like insufficient lubrication or overload using historical data received from wind turbines. The research prepared five neural network models based on historical data collected from wind turbines. The neural networks will help explain how the various input factors interact with the temperature of the generator bearings and hence predict failures. After that, two turbine instances were utilized to train the neural network models, and their efficiency was validated.

4.2 Process management

The advanced fault detection and diagnosis, recently discussed in the literature, significantly optimize complex industrial processes and ensure their efficacy. For example, inspired by their collaboration with (Liu et al., 2012), a novel reconstruction-based method for identifying and correcting faulty variables and blocks in continuous annealing processes is suggested. Separately, Hu et al. (2014) presented an adaptive fault detection and isolation method incorporating PCA and generalized likelihood ratio tests within the framework of the imperial smelting process. Compared with traditional SPC approaches, the PCA-based method enhances fault detection rates by approximately 22%. This illustrates its effectiveness in actual manufacturing settings. In their industrial case study on steel manufacturing (Jakubowski et al., 2024), the integration of PCA with a real-time sensor network generated an approximately 30% decrease in production downtime and enhanced

process stability, mitigating hidden anomalies. Such efforts manifest that refined diagnostics, as a technique, may considerably impact process reliability and performance.

4.3 Photovoltaic systems

The photovoltaic power generation systems are affected by some weather-related abnormalities in their external environment (Pei and Hao, 2019). Unfortunately, the nonlinear characteristics of solar applications make it difficult to classify and identify these anomalies. An innovative semi-supervised learning model based on graph theory was introduced in (Zhao et al., 2015), which is capable of effective hidden fault detection and classification in photovoltaic arrays using a limited amount of labelled data. The model showed a fault detection accuracy of 92.8% with a very small dataset, which is enough to make it a recommendation for industrial-scale solar farms. The AI-driven approach shows 30% less number of false alarms in comparison to the conventional threshold-based monitoring, making fault diagnosis more reliable. Combining findings by Pei and Hao (2019), critical fault parameters were determined, and voltage and current reference values were set for photovoltaic systems in establishing an appropriate Pearl detection scheme. The proposed method employed thresholds to determine a fault based on the same criteria and some existing thresholds to benchmark against these somewhat arbitrary reference thresholds, pinpointing actual malfunctions. The testing of their approach on PV arrays at an industrial scale reached a reported detection efficacy of 94.6%, well above the conventional threshold-based average of 83%. Notably, the real-time and AI-driven observation of anomalies reduced the panel degradation alert interval by 40%. It brought forth improved energy efficiency and longevity for the system. The analysis provided support to implement their proposed solution for practical applications.

4.4 Semiconductors

Some challenges in the semiconductor industry include formulating and implementing the high performance of FDC systems to recognize and categorize all the faults that occur during production. Real-time monitoring and FDD activities support the importance of ensuring safe and consistent production through controlling any possible departures from normal processes (Zhou et al., 2018). The overall semiconductor industry has accepted FDC as a key component of advanced process control solutions. Mainly, it focuses on production throughput and tool utilization rates, along with lowering cycle time (Wang et al., 2012). Tobacco has used several classical techniques to improve the processing inside the realm of FDC. That said, traditional FDD methods may not effectively identify the type and sources of defects during semiconductor wafer processing. Fan et al. (2020) found that RF methods have shown 18% less error than k-NN and PCA methods in wafer originating from semiconductor defects classification. Additionally, implementing FD-kNN in the industry led to 25% more isolated fault detection at semiconductor foundries, avoiding major wafer scrap costs and rework. AI-equipped FDD methods are said to increase

production yield by 12% across five semiconductor fabs, because of differing characteristics that older FDD styles would find hard to capture. The literature is full of evidence of advancements made in advanced FDD since the semiconductor industry last 3 decades (Susto et al., 2017).

He and Wang (2007) introduced the FD-kNN method, based on the k-NN approach, to address the shortcomings of existing PCA-based FDD approaches. The technique is dedicated to fault detection in semiconductor manufacturing processes characterized by nonlinearities and various operating modes. Random Forest (RF)-based methods were employed by (Shu Kai S. Fan et al., 2020) for processing time and steps in semiconductor manufacturing processes. Their method improved accuracy in fault classification by 15% regarding PCA-based techniques, showing that ensemble models capture complex fault patterns better. K-NN and naive Bayes classifiers were used in an ensemble model alongside the t-distributed stochastic neighbor embedding technique to visualize the emergent FDC parameters. Monitoring process conditions is important in systematically managing wind turbines and farms to incur lower operating costs. Based on an SVM-based FDI design, the wind turbine FDD was discussed in (Laouti et al., 2011). They dealt with windbreaks in a three-blade variable-speed horizontal-axis wind turbine backed by a full converter. Data-driven methodologies devoted to designing FDI for wind turbine systems regarding nonlinearities and unknown wind perturbations were examined in research (Li et al., 2023). Kusiak and Li (2011) proposed a system of fault prediction in wind turbines composed of three levels: problem identification, categorization by their criticalities, and predictions of various defects that SCADA systems and fault files will support.

4.5 Electric/smart vehicles

Reliability is an important feature of major components within the EV, including the battery pack, motors and sensors. Failure of these important components determines from which area the failure appears, to bring the vehicle back into performance and safety. Kang et al. (2023) have proposed an improved online multi-fault detection and diagnosis technique in EV battery packs. This research has focused on conceiving a capable system that behaves by simultaneously detects and diagnoses multiple faults in the battery system online. The proposed system is simultaneously performing two important operations: real-time monitoring of battery packs and deploying complex diagnostic algorithms capable of behavior identification amid fault patterns and prognostic modeling suited for battery health and capacity. This suggested system further improves reliability and battery safety for the EV systems. This new strategy in the battery pack diagnosis is solely devoted to the ideal fault identification focused towards understanding the operational capacity of the vehicle. This new methodology will now facilitate the development of the faulty signalling elements, erratic vehicle failures when there is no warning from the battery failure itself, which could cause stresses to the components and/or on the performance of the vehicle. Kang et al. (2023) reported that, in their large-scale testing of the battery health monitoring systems, their multi-fault detection model

enhanced fault isolation accuracy by 28% and provided a 15% reduction in premature battery replacements, signifying a major cost-saving to fleet administrators. In addition, the implementing of AI-centric battery diagnostic systems in the commercial pool of EVs has conferred a 23% mitigated risk of facing sudden battery failures, which enhances vehicle reliability and safety.

(Du, 2019) proposed fault detection of interior permanent-magnet synchronous motors (IPMSMs) used in electric drive applications for their interior permanent-magnet synchronous motor (IPMSM). IPMSMs are one of the types of motors used within electric vehicles. The in-vehicle faults of the motor caused by the inter-turn fault in an IPMSM can somewhat lower the performance level, leading to a total failure of the motor. High-frequency, low-resolution signals were suggested for specific purpose-based fault detection of an IPMSM in the case of an inter-turn fault, with this scheme being very suitable for the operation of electric vehicles. (Byun et al., 2019) reported a series of techniques for sensor fault detection and signal restoration, with an eye toward performance for intelligent vehicles in ADAS and FA applications. Regarding safe and effective performance and operational autonomous vehicles and technologies, reliable sensor information is paramount for both ADAS and full autonomous driving. This present work has developed schemes that successfully identify conditions of sensor fault and restore output signals as closely as possible to their original values, thereby reestablishing confidence in the signals for the functioning of safety-critical features. Their research addressed the sensor failure issue for the maintenance of performance effectiveness and the overall safety of intelligent vehicles.

4.6 Robotic systems

Fault detection and diagnosis (FDD) of robot systems is essential to ensuring operational safety and reliability. (Golombek et al. 2011) show the need for an online data-driven FDD approach in robot systems. Their work promotes online performance monitoring and adaptive diagnostics to detect and rapidly diagnose faults when they occur, which may be needed in dynamic and uncertain environments. Experimental testing on an industrial robot arm showed that data-driven FDD models shortened fault response time by 45% compared to the conventional rule-based method. Real-time adaptive diagnostic use in robotic manufacturing lines was shown by (Golombek et al. 2011) to boost uptime by 30% and prevent unwarranted shutdowns while maximally enhancing running efficiency. Data-driven FDD uses large datasets to identify aberrant behavior and enable early intervention, which can assist systems in minimizing downtime and maximizing system overall performance.

On the other hand, (Cho et al. 2018) employ a model-based approach for fault diagnosis of robotic systems. The model-based FDD technique can leverage pre-existing models of system behaviour to recognize deviations and diagnostic faults. Their approach allows comparing real system behaviours with existing models to generate a diagnostic rationale. Model-based FDD prioritizes formalized theoretical knowledge accompanied by empirical knowledge to construct a strong framework of

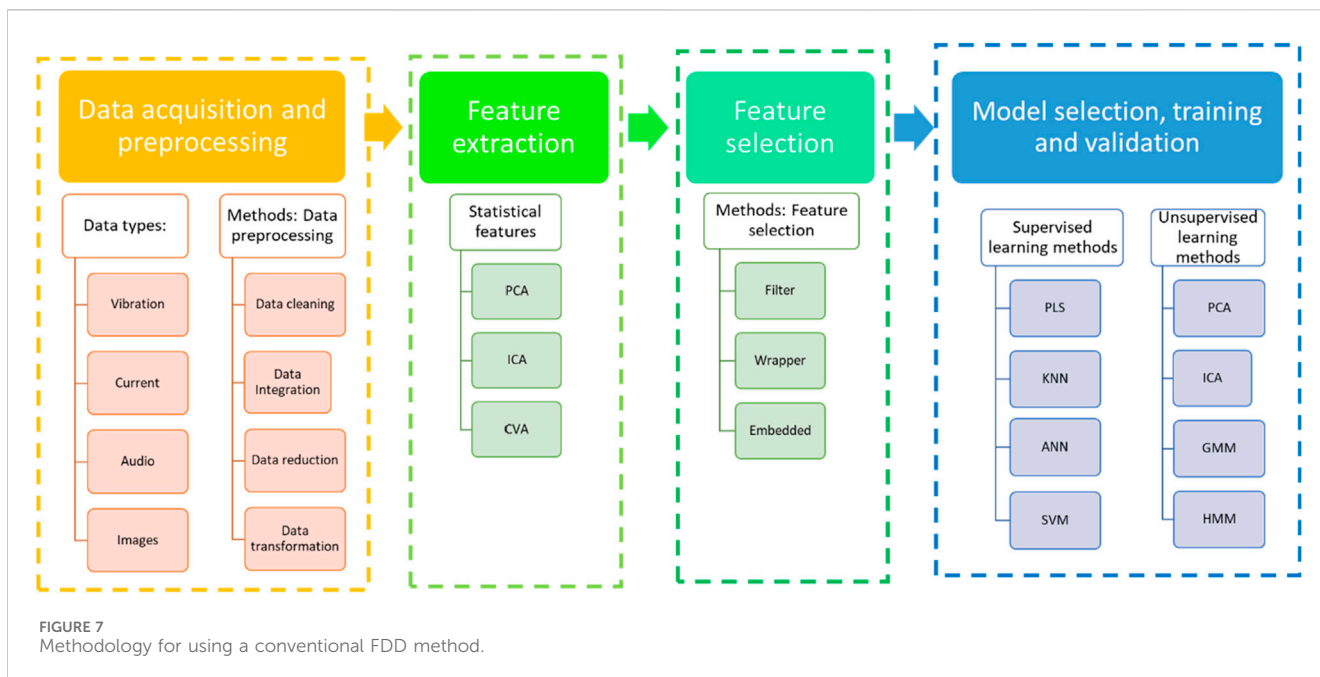
diagnosis that bestows greater precision and dependability to robot systems.

4.7 Building system

Fault detection and diagnosis (FDD) of a building system has many approaches depending on the component and operational option. (Heracleous et al. 2016) researched critical infrastructure (specifically water supply systems) with special emphasis on online fault detection approaches to track integrity and reliability. The authors noted how diagnostics provide real-time tracking of these services, guaranteeing continuity of operations. (Tran et al. 2015) assessed a strong online fault detection and diagnosis method for centrifugal chiller systems for building energy efficiency. The authors correlated fault detection and energy management procedures in their methodology to demonstrate how diagnostics methods can improve efficiency in operation and sustainability. (Kim and Katipamula 2018) provide a comprehensive review of FDD methodologies applied in building systems, detailing different methods and applications in building components. Their review points to the contribution of FDD methodologies to building performance, giving an insight into future trends. (Kumar et al. 2021) describe a sensor-based building fault detection and diagnostic system in terms of developing new types of sensor networks and their ability to enhance diagnostics. Their research illustrates how sensors used in compositions can enable a better ability for diagnosis.

(Kim and Lee 2021) established the establishment of virtual sensors to detect faults in a room air conditioner, establishing a more precise diagnostics process and an outcome of fewer maintenance operations. This new use utilized a virtual technology process because it considers fault detection of building systems. Lastly et al. (2006) outlined a hierarchical rule-based approach to assessing faults and fault diagnosis in HVAC systems. In context, evidence emerged that a rule-based definition of faults can cope with the complexity of HVAC diagnostic machines. Expert FDD approaches in HVAC systems, for instance, by (Schein and Bushby 2006), enhanced the fault detection rate by 35% compared to inspection-based manual assessment. A large-scale deployment of a commercial building (Kumar et al., 2021) showed 20% less energy wastage with early identification of system poor performance, leading to enormous cost reductions. Their rule-based method is one example of a systematic diagnosis method for reliability to make buildings efficient.

In short, FDD is a core function to ensure manufacturing systems are reliable, efficient, and safe in today's sector. This system diagnosis and fault detection through a systematic approach enables a head start on early detection, which helps reduce downtime and costly disturbance by an FDD system. Advanced algorithms and data analysis run these systems to monitor equipment performance constantly for predicting oncoming failures and practical insights into timely maintenance. As more sectors focus on digital and automated technologies, integrating FDD ensures that operations remain optimized to reduce operating expenses and extend the lifespan of main assets. Ultimately, this pays dividends in maintaining the integrity of manufacturing and keeps industries competitive and sustainable



overall. Comparative evaluation of various FDD techniques across various industries demonstrates significant trade-offs: AI techniques (ANN, SVM, and RF) possess high accuracy (typically >90%) but require extensive training data and computational resources. Conversely, model-based techniques such as PCA and k-NN are interpretable but less effective in handling nonlinear system dynamics. Industrial case studies demonstrate how AI-enhanced FDD systems can reduce unplanned downtime by up to 40%, significantly improving asset reliability and cost savings. Such results underscore the revolutionary potential of advanced fault detection methods in real-world applications.

5 Fault detection methods

Maintaining optimal performance in complex industrial processes with multiple faults can be challenging. To ensure optimal process supervision, fault detection and isolation are two common approaches. Among these, FDD is a highly significant technique for controlling process performance, as most industries strive to improve their operations through increased FDD (Hu et al., 2014). The primary objectives of FDD are to monitor the current state of a process and identify any flaws, including their characteristics and underlying causes (Pei and Hao, 2019). Process monitoring serves various purposes, the most crucial being the prevention of interruptions. Effective management of potential issues during the process is the primary goal of process supervision. Therefore, FDD plays a vital role in numerous industrial processes and remains a key research focus due to its significance in ensuring safe, productive, and effective operations (Shu Kai S Fan et al., 2020). Figure 7 illustrates the four phases of deploying a conventional FDD method. To ensure an accurate diagnosis, it is crucial to correctly identify faults, which involves determining the type of problem, its severity, location, time of

detection, and behaviour. Combining fault isolation and identification is often known as the fault diagnostic step (Kusiak and Li, 2011). Nonetheless, fault diagnosis typically encompasses a broader range of tasks, including fault detection, isolation, identification, classification, and evaluation.

Numerous FDD methods have been developed and implemented across various industries over the past 3 decades. Researchers and practitioners have proposed various FDD methods, including model-based, observer-based methods (Bouchikhi et al., 2015) and data-driven methods, such as PCA and NN. In this regard, we will delve deeper into different approaches to FDD.

5.1 Data-driven techniques

Data-driven FDD techniques have gained widespread attention across various industrial sectors due to their effectiveness in monitoring complex processes (Ardali et al., 2024). Table 2 presents a classification based on the system's features to provide an overview of data-driven FDD techniques.

The following sections review some of the most common and powerful methods adopted in FDD for various industrial applications: PCA, ICA, SVM, kNN, and NN. All these techniques have different powers and capabilities regarding fault identification, classification, and diagnosis, which makes them an asset in current industrial monitoring systems.

5.1.1 Principal component analysis

Principal Component Analysis (PCA) is a statistical technique used to reduce the dimensionality of data while maintaining its disparity. In the context of FDD, PCA is employed to transform a large set of process variables into a smaller set of uncorrelated variables called principal components (PCs) (Mansouri et al., 2016). Typically, a dataset is collected from machine sensors under normal

TABLE 2 Data-driven methods.

System	Methods	Advantages	Disadvantages
Dynamic Systems	HMM(Galagedarage Don and Khan, 2019), Dissimilarity-based PLS (Harmouche et al., 2015) Dynamic NN (Sedej et al., 2022)	- Good at capturing temporal dependencies and dynamic behavior - Effective for online monitoring and control	- Requires a large amount of data - Complexity in parameter tuning and model selection
Non-linear Systems	Kernel PCA, Kernel ICA (Choi et al., 2005), ANN ((Seid Ahmed, 2024)), Kernel PLS (Botre et al., 2016), KNN (Guo et al., 2003)	- Capable of modelling complex, non-linear relationships - High flexibility in capturing system behavior	- Computationally intensive - Prone to overfitting with limited data - Interpretability can be challenging
Non-Gaussian Systems	ICA (Fan and Wang, 2014) Gaussian Mixture Model (Tong et al., 2018), Kernel PLS (Harmouche et al., 2015), Bayesian Estimation (Gonzalez et al., 2015)	- Suitable for non-Gaussian data distributions - Good at uncovering underlying independent components	- Sensitive to noise and outliers - May require complex pre-processing and tuning
Time-varying Systems	Recursive PCA (Sun and Jia, 2019), Isolation Estimation (Park et al., 2020)	- Adaptive to changes over time - Effective for online updating of models	- May have difficulty with abrupt changes - Can be sensitive to initialization and parameter settings
Non-Stationary Systems	SVM (Gnanamalar et al., 2023), Hidden Semi-Markov Model (Bighamian et al., 2014), HMM (Gonzalez et al., 2015), Monte Carlo Simulation SVM (Gnanamalar et al., 2023)	- Capable of handling data with varying statistical properties - Suitable for long-term system monitoring	- High computational cost - Complexity in handling large-scale data - Model interpretability issues

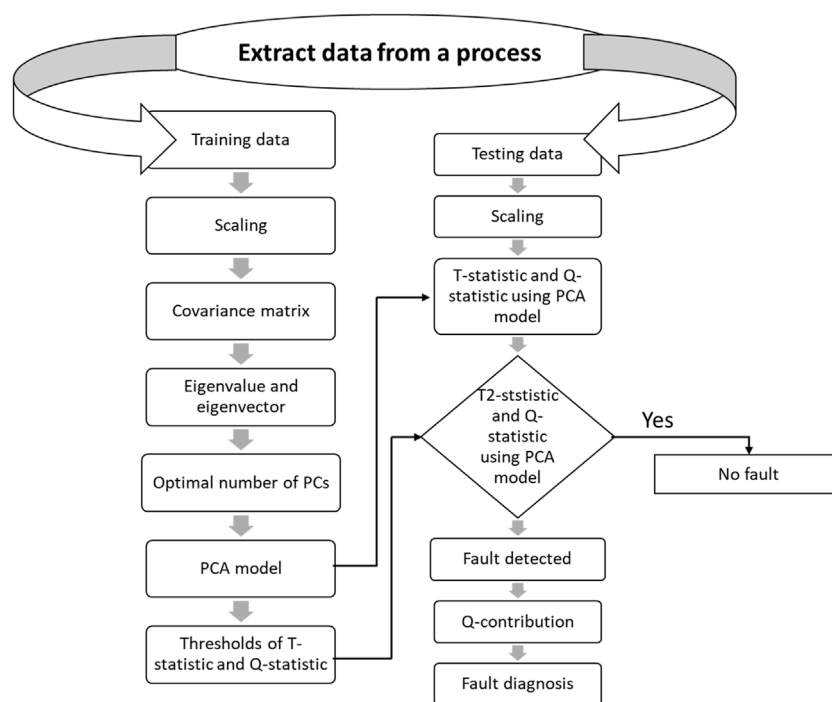
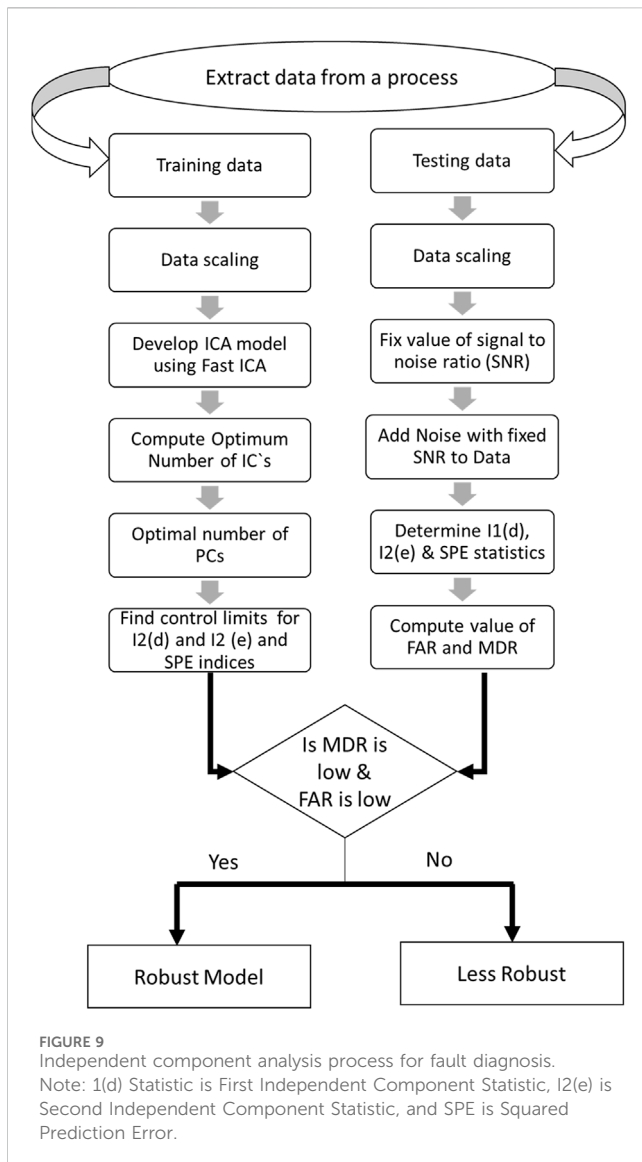


FIGURE 8 Principal component analysis process for fault diagnosis.

operating conditions. Then, the data is scaled, and the covariance/eigenvectors matrix is computed. The original data is projected onto the principal components, reducing dimensionality, before applying statistical process control techniques to the principal components to detect deviations or anomalies in machine data and identify the potential root causes of the fault. Figure 8 shows the fault diagnosis process using PCA. Some advantages of PCA-based FDD methods

include dimensionality reduction, noise reduction, fault detection sensitivity, and interpretability. While its limitations are: linearity, sensitivity to outliers, and difficulty in handling nonlinear processes. Despite these limitations, PCA remains a valuable tool in FDD due to its simplicity and effectiveness in many industrial applications. Some works in which PCA-based methods are used for FDD are discussed as follows.

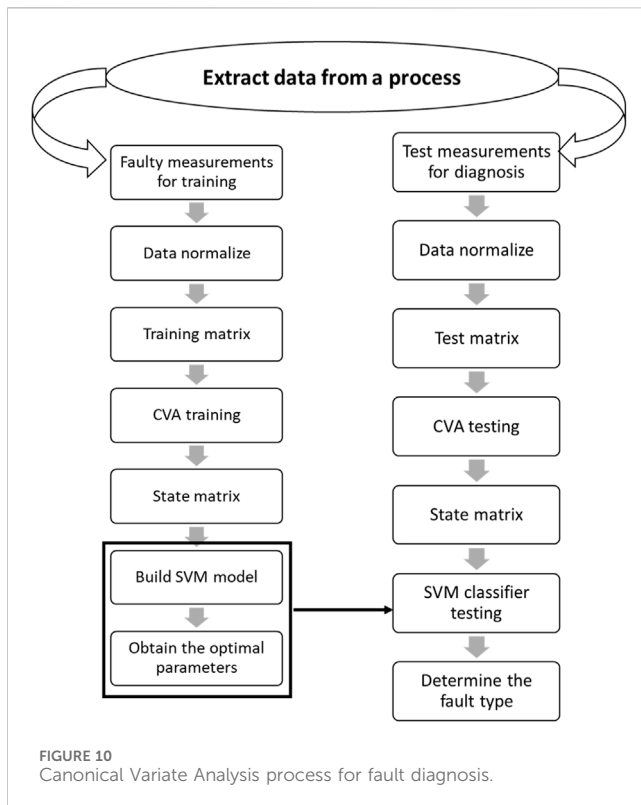


Sun and Jia (2019) recommend a novel FDD technique based on PCA. They developed a new method of feature engineering to achieve early fault prediction and lower false alarm rates simultaneously. Implementing the proposed technique in a boiler leak detection case proved its effectiveness in spotting the area where the problems are. Harrou et al. (2013) showed that PCA-based techniques effectively detect unknown faults in complex industrial processes, among other statistical FDD approaches. Dynamic principal component analysis and reconstruction-based contribution techniques are two PCA types widely used in FDD implementations. To maximize the reduction of combined index (RCI), Liu et al. (2012) suggested a contribution plot method based on a missing data approach. Liu et al. (2022) used a multi-level PCA. Harrou et al. (2013) used it on data from a simulated continuously stirred tank reactor (CSTR) to increase the reliability of fault detection without a process model. Lau et al. (2013) presented a framework for detecting issues in real-time manufacturing processes using multiscale principal component analysis (MSPCA) and ANFIS to infer fault-signal correlation from process data.

Estimating and detecting minor and random variations due to initial faults was done using the pullback-Liebler divergence (KLD) theory by (Harmouche et al., 2015). This penetration testing approach suggested that the fault-to-noise ratio (FNR) should be used as an index to compare the gravity of faults and the surrounding noise. Dynamic principal components analysis and the Bayesian inference (DPCA-BI) is a method for real-time process monitoring proposed by Huang and Yan (2015), which carries out the Jarque-Bera test to determine whether a Gaussian or a non-Gaussian distribution exists. In a recent study by Mnassri et al. (2015), the RBC method and RBC ratio (RBCR) have been theorized as advancements in data-driven fault diagnostics, which can predict and fix major multi-dimensional faults. Bakdi et al. (2017) discussed a novel data-driven FDD method using PCA to model the cement manufacturing process and adaptive thresholds to assess two frequently employed multivariate statistics. They used a modified exponentially weighted moving average (EWMA) chart to adjust the threshold numbers dynamically. The effectiveness of the suggested FDD scheme was verified by testing it on some error scenarios. Several obstacles, described in (Du, 2019), make detecting problems in actual manufacturing processes difficult. Feature extraction is widely acknowledged as a critical step towards effective FDD (Lau et al., 2013). Du (2019) suggested a feature extraction method using bispectral features and a one-class classifier to characterize process dynamics and address class imbalance problems.

5.1.2 Independent component analysis

Independent component analysis (ICA) is a statistical technique to identify underlying factors or components in multivariate data. Unlike PCA, which seeks orthogonal components, ICA seeks statistically independent components. ICA can be a powerful tool for extracting hidden information from sensor data to uncover underlying physical processes, fault modes, or the root cause of the fault. The fault diagnosis process using ICA is shown in Figure 9. ICA is useful in feature extraction, noise reduction, and fault isolation. However, it has several limitations, including non-Gaussianity, computational intensity, etc. The use of FDD techniques employing PCA can quickly identify and resolve issues. However, PCA's ability to detect complex linear or nonlinear relationships among process variables may be limited because of its Gaussian distribution assumption. To deal with non-Gaussian and nonlinear industrial processes, Fan and Wang (2014) developed a kernel dynamic independent component analysis method, which leverages the auto regressive (AR) model and ICA and kernel methods to extract dynamic characteristics. Nonlinear contribution plots were also created to aid in fault detection. Tong et al. (2018) proposed the double-layer ensemble monitoring technique based on a modified version of ICA called DEMICA to tackle non-Gaussian process monitoring challenges. The DEMICA approach used a double-layer Bayesian inference method for tracking indices and included all feasible MICA models as an ensemble. Traditional FDD methods typically utilize principal component analysis, which assumes a normal distribution of process data and statistical independence between samples. However, these assumptions often result in many false alarms and issues with missing detections in process monitoring.



5.1.3 Canonical Variate Analysis

Canonical variate analysis (CVA) is another statistical technique used to find linear combinations of two sets of variables that have the maximum correlation with each other (Jiang et al., 2015). CVA is used to monitor the projection of new data onto the canonical variate space and analyze the contributions of the original variables to the canonical variates to identify the root cause of the fault. As a result, CVA can reduce the dimensionality of the data, making it computationally efficient, and can effectively discriminate between different fault modes based on the signal associated with multiple process variables. Figure 10 illustrates the fault diagnosis process using CVA. Nevertheless, CVA assumes linear relationships between variables, which might not be suitable for all systems. Moreover, outliers can significantly impact the results of CVA; a data-driven multivariate method would perform well when there are serial correlations among process variables.

Juricek et al. (2004) proposed a subspace identification algorithm based on the CVA approach. Russell et al. (2000) also integrated the CVA method with dynamic principal component analysis to realize a residual generator for fault detection systems. Jiang et al. (2015) applied CVA to derive the retained states and residual space and then, respectively, established fault detection and identification schemes based on these two spaces.

5.1.4 Support vector machines

Support vector machines (SVMs) are powerful machine learning algorithms that have recently found widespread application (Ding et al., 2016). They excel at classification tasks, making them ideal for distinguishing between normal and faulty system states. SVMs often achieve high accuracy in separating different fault classes, even for cases involving complex datasets that exhibit numerous features.

Moreover, SVMs are relatively insensitive to outliers; hence, they can be effectively applied to classification and regression problems. However, SVMs can be computationally expensive for large datasets. Moreover, choosing the appropriate kernel function (linear, polynomial, radial basis function, etc.) is crucial for SVM performance. Support vector machines (SVMs) have been widely utilized in process monitoring and fault diagnosis applications to identify anomalous patterns and classify them based on fault type. Figure 11 illustrates the fault diagnosis process using SVMs.

Widodo and Yang (2007) conducted a comprehensive survey of SVM methods for machine health tracking and fault diagnosis, highlighting future directions in expertise and problem-solving. Another research (Park et al., 2011) proposed a novel fault extraction and identification approach using cubic spline regression and Support Vector Machines (SVMs). They used SVMs to build a classifier and cubic spline regression to determine transition points between discrete steps. Banerjee et al. (2010) tackled the dynamic motor condition fault diagnosis issue by employing SVMs and the short-time Fourier transform to classify the fault signals. Bordoloi and Tiwari (2014) suggested an SVM method for multi-fault classification of machine gears based on frequency domain information, using grid search, genetic algorithm, and artificial bee colony algorithm to determine the SVM model's optimal parameters. Xiao et al. (2018) proposed a procedure to determine the optimum parameters of the one-class SVM for multiclass classification. Jing and Hou (2015) compared PCA and SVM for fault categorization in a multiple-class problem.

5.1.5 K-Nearest Neighbor

K-Nearest Neighbors (KNN) is a non-parametric machine learning algorithm that classifies new data points based on the majority class of their k-nearest neighbours in the training dataset. In FDD, KNN can classify system states as normal or faulty by extracting relevant features from normal and faulty system operation data. KNN is relatively easy to understand and implement. It can also capture complex relationships between features and class labels. Nevertheless, KNN can be computationally expensive for large datasets and is usually sensitive to noise in the data, as it relies on distance calculations; thus, KNN outcomes might result in outcomes with bias. Figure 12 illustrates the fault diagnosis process using KNN is illustrated in Figure 12. Various approaches based on KNN effectively detect unique characteristics of real-world industrial processes, such as multimode batch trajectories and non-linearity (Kong et al., 2019). However, conventional multivariate methods can be challenging to use in processes with non-Gaussian distributions or correlations among many process factors, such as semiconductor manufacturing. To address these issues, researchers have proposed novel techniques, such as an adaptive Mahalanobis distance-based KNN detector (k-NND) (Verdier and Ferreira, 2011) and a semi-supervised version of the Fisher discriminant analysis (SFDA) model that considers both labelled and unlabeled data. Class imbalance is another common challenge in fault detection. Researchers have proposed methods such as incremental clustering-based fault detection (IC-FDM) (Kwak et al., 2015) and a hybrid fault detection technique that combines random projection with KNN (RPkNN) (Zhou et al., 2015). Novel distance-based fault detection

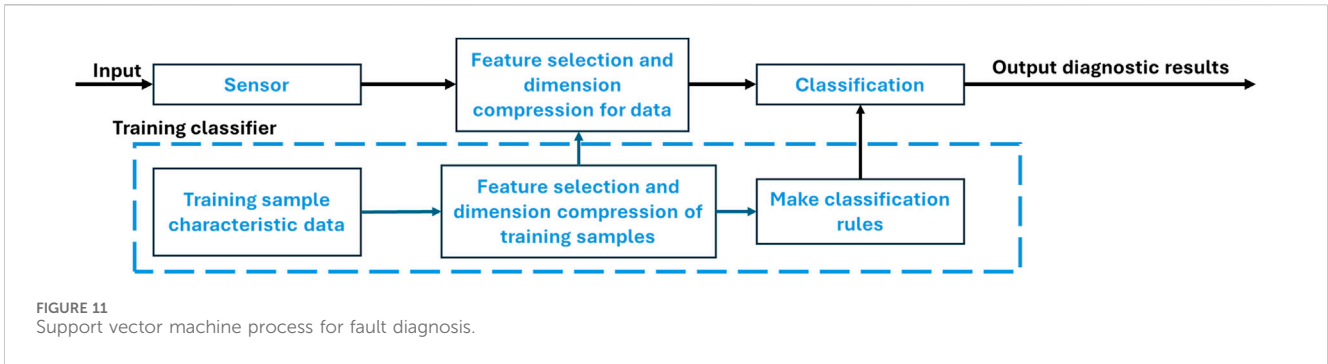


FIGURE 11 Support vector machine process for fault diagnosis.

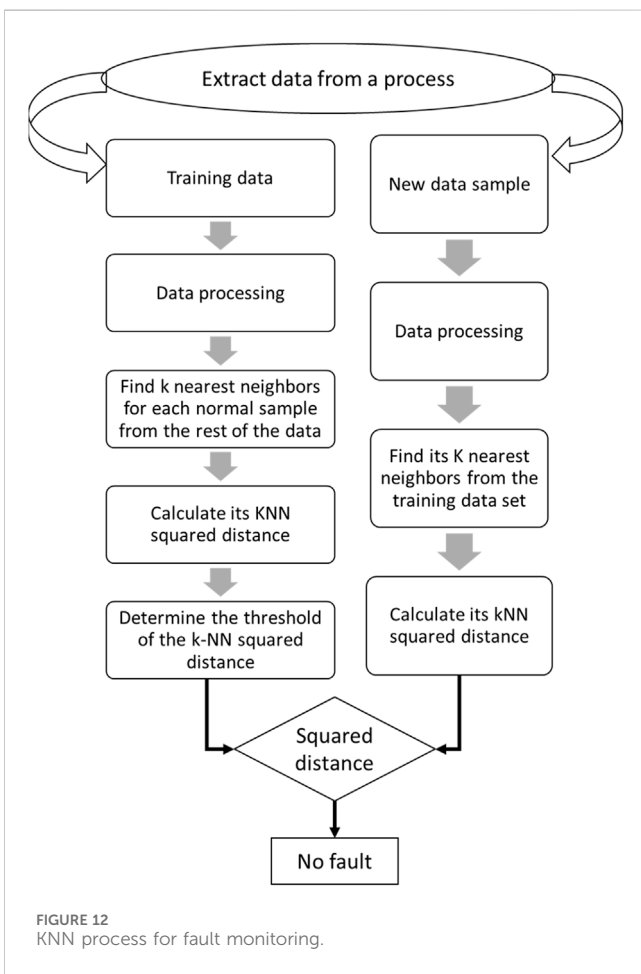


FIGURE 12 KNN process for fault monitoring.

methods, such as the KNN weighted distance method (Zhang et al., 2019), have also been suggested, incorporating statistics designed to improve fault detection.

5.1.6 Neural network

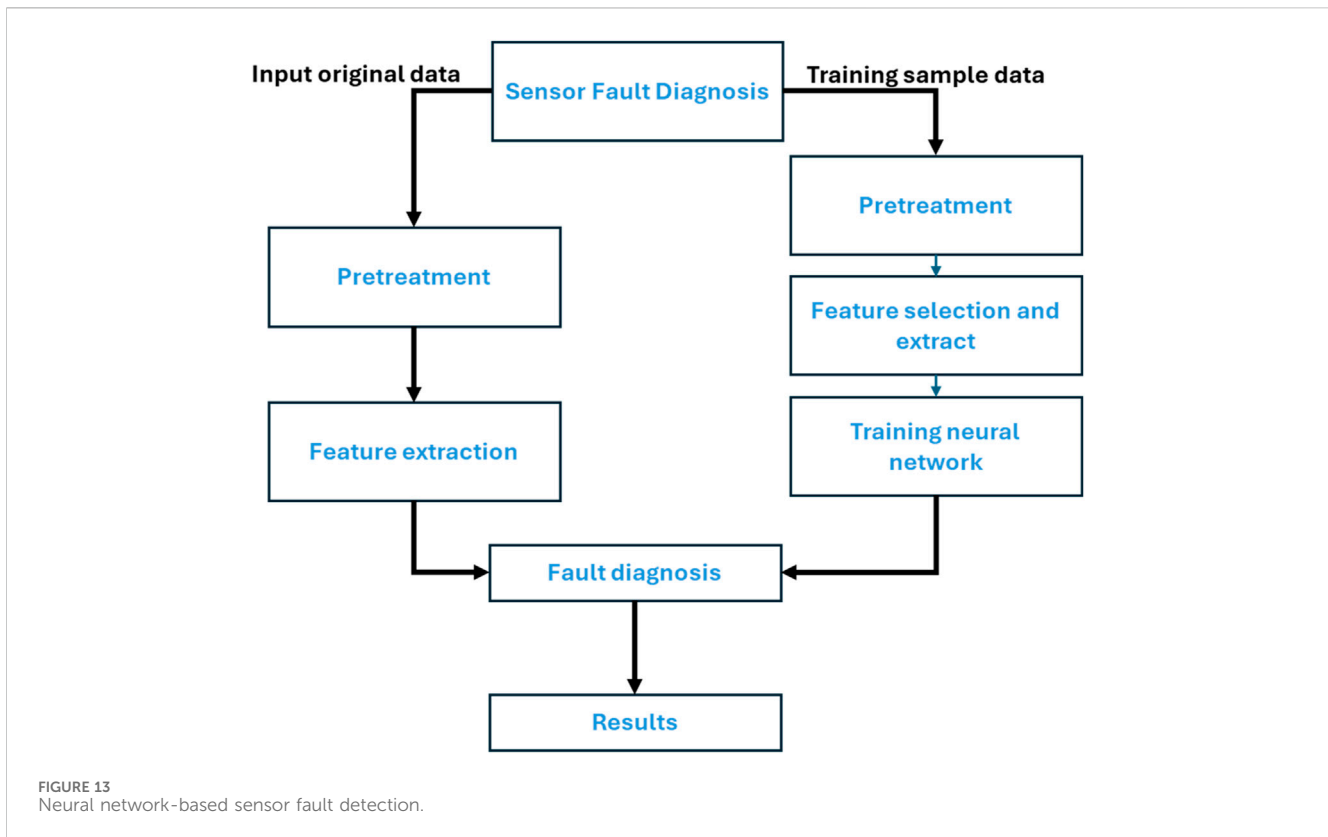
Neural networks are powerful machine-learning models inspired by the human brain (Seid Ahmed, 2024). They are particularly effective in handling complex patterns and nonlinear relationships within data, making them well-suited for FDD. Neural networks can achieve high accuracy in FDD, automatically learning relevant features from machine data. However, training and deploying large neural networks can be computationally expensive. Furthermore, neural networks are prone to overfitting

and often require extensive data for training. Several methods for FDD in manufacturing processes using intelligent systems were discussed earlier in section 3. One such approach involves using a neural network (NN) to detect early signs of machine failure (Seid Ahmed, 2024).

Due to neural networks' capabilities in converting inputs into non-linear outputs, they can be utilized for diagnosing sensor faults. Additionally, they can handle issues related to speculation and memory (Ren and Lv, 2014). When implementing a neural network for sensor fault detection, the selection of network architecture and activation functions depends on the output vector's dimensionality and the number of faults to be identified. The network is trained using data collected from fault samples to ensure accurate fault detection and classification. The network is fine-tuned with an inertia factor, learning rate, and other data-based parameters to achieve highly accurate results. In Figure 13, the flowchart of the detection procedure is depicted. The general concepts behind most typical neural network applications have been well-established for several years, indicating that the field is relatively mature. Presently, research on diagnostic methods using neural networks mainly focuses on enhancing the algorithm for training data (Seid Ahmed, 2024).

Using neural networks in sensors to diagnose faults in complex systems is common. Research (Ren and Lv, 2014) integrated an artificial neural network into the combined harvester to detect sensor failures. However, the inherent static characteristic of the data used usually leads to extended time to isolate the fault. In another research (Ferentinos et al., 2003), the weights were trained with backpropagation. The authors could reduce the effects of plant imperfections on training precision by adopting a classical training method that eliminates characteristics of information from specific sensors. Based on their findings, they concluded that it is possible to achieve online fault detection with the help of neural networks. The unmanned aerial vehicles, also known as UAVs, are a new and rapidly evolving market. Considering the emulation of the permanent breakdown of sensors in UAVs, Samy et al. (2011) proposed using the EMRAN-NN to emulate the permanent breakdown of sensors in UAVs based on the current neural network (NN) scheme. They then tested the feasibility of this method using collected data. However, during the Fault-Tolerant Control System (FTCS) demonstration, they observed some above-average Fault Activation Events (FAEs) that seemed significantly higher than normal and should be further investigated.

A study (Habriansyah and Andani and Zainuddin, 2017) found that maintaining strict control over the short-term storage



temperature of flour is crucial. The study introduced a technique for detecting temperature sensor faults based on online learning. This technique utilizes a feedback artificial neural network trained with a random sequence-learning algorithm to identify abnormal sensor readings. Additionally, auxiliary sensors were used to replace malfunctioning main sensors through redundant analysis, which improved system efficiency. Previous research indicates that neural networks are advantageous for real-time operation and direct processing of time sequence data, making them a widely used methodology. However, neural networks also have inherent flaws, requiring numerous training samples and showing weak generalization capabilities, which could pose challenges in fault detection efforts.

5.2 Model-based techniques

Due to the limitations of traditional FDD methods, various alternative approaches based on models, particularly process models or mathematical models, have been investigated. Model-based FDD methods utilize the interdependencies among multiple process variables. Therefore, users must comprehend the process model before implementing this technique. Pourbabaee et al. (2016) discussed various model-based approaches to fault detection and isolation (FDI) for automated systems. The aim was to enhance the robustness of the approach to modelling errors using state estimation and parameter identification strategies (Cho et al., 2018). Also focused on model-based approaches while discussing fault detection and isolation techniques. In addition, statistical residual testing techniques were discussed for detecting abrupt

faults and the methods of implementing a reconfigurable control strategy based on a model. With their innovative combined FDI scheme, Shahid et al. (2023) used subspace model identification (SMI) to create a precise model by keeping tabs on the individual stages of the process, and the authors' proposed scheme was used to produce useful statistics from the multivariate residual. By implementing three commercial applications with sensor faults, process faults, and actuator faults, they proved the efficacy of the suggested scheme.

Azmi and Yazdizadeh (2022) developed new strategies for detecting system faults using observer-based diagnostic techniques. They extensively analyzed various observer-based approaches for FDD. Wu et al. (2017) proposed an advanced approach to FDD for nonlinear systems with multiple early-stage sensor failures in a research paper. The strategy involves a sequential combination of two techniques -total measurable fault information-based residual (ToMFIR) and sliding mode observer (SMO) - for measurement repair. The original systems were divided into two sections using a state and output transformation method to make the fault diagnosis more accurate. Piltan and Kim (2018) introduced a variable structure feedback linearization observer (FLO) with FDD design in rotating machinery. Bernardi and Adam (2020) proposed two observer types for the FDD approach, using a linear parameter-varying system for nonlinear chemical processes.

A survey by Li et al. (2022) identified improvements in multiple monitoring and fault diagnosis techniques that could be incorporated. Yin et al. (2022) presented a solution to FDI by introducing a new strategy based on the Parity Space Optimization concept. By transposing the converted parity relations, they created an optimum transformation matrix. They

TABLE 3 Hybrid fault diagnosis techniques.

Techniques	Approach 1	Approach 2
Multiple-Model Banks and Spectral Analysis (Fathi and Bolandi, 2024)	Failure detection using Recursive Least Squares (RLS) with a hybrid threshold that adapts in real time	Fault severity index calculation based on typical fault frequencies
Evaluation Model and Real-Time Data (Hekmat and Ravanmehr, 2016)	Uses analog signals like temperature, current, and voltage, while bus network communications are digital	Wiring harness temperature assessment model for real-time monitoring
Deep Learning and Signal Processing (Shao et al., 2017)	Multilayer Perceptron (MLP) for domain knowledge generation	Convolutional Neural Networks (CNN) for feature extraction and fault classification
Machine Learning and Statistical Analysis (Kankar et al., 2011)	Initial fault detection threshold based on permutation entropy and Chebyshev's theorem	Variational Mode Decomposition (VMD) with hybrid features for fault identification and severity assessment
Data Mining and Supervised Learning (Zhao et al., 2015)	Stacked Autoencoder for labeling unlabeled historical process data	Clustering using t-Distributed Stochastic Neighbor Embedding (t-SNE) for fault classification

applied the pseudo-inverse and actuator estimation methods and tried to model the process with a cumulative sum chart, which helps to register when the residual variance changes. The method was tested in simulation and real fault cases of aircraft control surface actuators to substantiate the proposed solution.

5.3 Knowledge-based techniques

FDD techniques based on multivariate statistics typically analyze process failure or anomaly data. When creating or employing FDD strategies for a wide range of industrial processes, knowledge or principles should be considered in addition to the process data. Knowledge-based approaches were developed to address this issue, benefiting several complex manufacturing processes (Pourbabaee et al., 2016). Representative knowledge-based FDD techniques include the fault-model-based cause-and-effect analysis approach and human-reasoning-based expert systems (Khireddine et al., 2014). Lou et al. (2022) proposed an error-detection system for chemical processes, incorporating Multivariate Statistical Process Monitoring (MSPM) into an existing knowledge-based FD framework. Peng et al. (2018) employed a fuzzy neural network approach for fault diagnosis in nuclear power stations. Meanwhile, Gonzalez et al. (2015) utilized process knowledge within a Bayesian network (BN) method for data compression to enhance system clarity. Chen et al. (2022) proposed a BN model in multiple model architectures, considering previous process knowledge and process data, to detect failures in HVAC systems. Don and Khan (2019) introduced a unified FDD approach utilizing a Hidden Markov Model (HMM) for fault detection and a Bayesian Network (BN) for root cause analysis. This integrated method outperforms purely data-driven approaches, leveraging the log-likelihood method for BN conditional probabilities and the input junction tree algorithm.

5.4 Hybrid-based techniques

Hybrid FDD approaches show great potential for practical use both in the short and long term. The most recent studies on hybrid methods are listed in Table 3. Table 3 summarizes the most recent

hybrid FDD techniques, illustrating the various approaches combining model-based methods with signal processing and machine learning techniques. The listed studies highlight the diversity in hybrid methodologies and their potential for more accurate and adaptive fault detection.

These approaches use a combination of model-based and signal-processing techniques to develop more robust designs for FDD. For instance, researchers have employed multiple model banks and multi-stage filtering schemes (Fathi and Bolandi, 2024) that utilize parameter and state estimation techniques. Signal processing techniques such as statistical and spectral analysis alongside deep learning have also been employed (Qian et al., 2022). The development of automatic hybrid FDD schemes presents significant challenges, such as designing a unified framework and ensuring the smooth interaction of different classifiers at different levels of abstraction. As the complexity of systems increases, these issues are becoming increasingly important. Another significant challenge in developing hybrid FDD schemes is ensuring the seamless integration of model-based approaches with signal-processing techniques. Data compatibility across different levels of abstraction, ranging from low-level sensor signals to high-level system states, must be addressed. Furthermore, managing the computational demands of real-time FDD, especially when deep learning models are integrated, presents a crucial hurdle. The success of hybrid FDD will rely on the robustness of its interaction mechanism, which must address compatibility, conflict resolution, and scalability.

Hybrid techniques often rely on effective fusion mechanisms to combine the results of different models. For instance, decision fusion can be applied where multiple classifiers' outputs are integrated based on a majority voting system or weighted averages. In cases where models conflict, strategies like priority-based resolution or consensus models may be employed to ensure accurate fault diagnosis. As complex systems grow, these mechanisms must be adaptable, scalable, and robust to handle varying data conditions.

5.5 Comparative analysis

Table 3 presents a comparative analysis of four Fault Detection and Diagnosis (FDD) techniques: data-driven, model-based, knowledge-based, and hybrid-based approaches based on the thorough review above. It can be concluded that:

TABLE 4 Comparative analysis of FDD techniques.

Aspect	Data-driven FDD	Model-based FDD	Knowledge-based FDD	Hybrid-based FDD
Methodology	Utilizes historical data with statistical and machine learning techniques	Relies on mathematical or physics-based models to describe system behavior	Based on expert-defined rules, heuristics, and domain knowledge	Integrates data-driven, model-based, and expert knowledge approaches
Computational Complexity	Moderate to high; complexity increases with large datasets	High; requires complex simulations and computations	Low to moderate; depends on rule scope and knowledge base	High; combines models, large datasets, and rule-based reasoning
Industry Applications	Manufacturing, HVAC, aerospace, automotive	Aerospace, automotive, power systems, industrial automation	Medicine, safety-critical systems, expert control systems	Robotics, advanced manufacturing, cyber-physical systems
Performance	Effective with high-quality data but sensitive to noise and missing data	Highly accurate when models align with real-world dynamics	Reliable if rule sets are complete and well-defined	High performance, often surpassing individual approaches when well-optimized
Advantages	Adapts to diverse datasets, requiring minimal system knowledge	Provides accurate fault detection for predictable system behavior	Fast and interpretable results based on expert-defined rules	Flexible, robust, and capable of handling complex systems
Limitations	Dependent on data quality, struggles with sparse or noisy data	Requires expensive and time-consuming model development	Performance is limited by expert knowledge and predefined rules	High implementation cost and complexity
Common Techniques	PCA, neural networks, clustering, deep learning	Observers, Kalman filters, and redundancy checks	Fuzzy logic, rule-based systems, expert systems	Bayesian networks, ensemble methods, hybrid AI models

- Data-driven FDD leverages historical data and employs statistical or machine-learning methods to identify patterns and detect faults. This approach is highly adaptable and practical in environments with extensive and high-quality data. However, its reliance on large datasets can be a limitation in scenarios with limited or noisy data.
- Model-based FDD relies on physical or mathematical models of the system for normal behavior prediction and fault reflecting deviation detection. High accuracy is provided by this method if the system dynamics are well understood and well modeled. The main drawback is the huge computational effort required to develop and maintain these models, especially for complex systems.
- Knowledge-based FDD employs expert knowledge in the form of pre-defined rules and heuristics to detect and diagnose faults. This approach is advantageous regarding interpretability and speed and can be applied to applications for which expert knowledge is readily available. Its performance is restricted by the completeness and accuracy of the expert knowledge embedded in the system.
- Hybrid-based FDD integrates data-driven, model-based, and knowledge-based approaches to take advantage of their strengths. The hybrid strategy enhances flexibility and robustness, particularly for complex systems where a single approach may be insufficient. Nevertheless, the integration can be complex and challenging.

Each FDD technique offers unique benefits and limitations, making them suitable for different industrial applications, as detailed in Table 4. Data-driven methods excel in data-rich environments, model-based approaches are ideal for systems with well-defined dynamics, knowledge-based techniques provide quick and understandable diagnostics, and hybrid methods offer a balanced solution for complex and multifaceted systems.

This comparison underscores the importance of selecting an FDD technique that aligns with the specific requirements and

constraints of the application. While more complex, hybrid-based approaches often provide the most robust and versatile solutions for sophisticated fault detection and diagnosis needs.

5.6 Implementation of fault detection and diagnosis techniques in industrial settings

Implementing FDD methods in industrial environments requires systematic integration of the entire chain from data acquisition, processing, and model selection to training, real-time observation, and continual improvement. Each of these resources contributes essential importance to realizing that the FDD system has to be trustworthy, reliable, accurate, and able to respond and work in an intricate context of various industrial domains with very few iterations on behalf of human execution (Saeed et al., 2025).

The data acquisition is the first step in implementing the FDD system; it consists of the real-time operational data collection from several sensors. Various sensors are deployed in industrial systems to monitor key parameters such as temperature, pressure, vibration, current, voltage, and flow rate. The sensor selection depends on the fault type that must be detected. For instance, vibration sensors are very important in rotating machinery, while thermal analysis requires infrared sensors in electrical systems (Vishwakarma et al., 2017). It is necessary to ensure that high-resolution and high-frequency data acquisition systems are employed to catch the transient fault conditions that might be unnoticed in low-sampling-rate data. The processed raw data is generally stored in industrial databases, such as Supervisory Control and Data Acquisition systems, edge devices, or cloud-based infrastructures (Surucu et al., 2023).

After data collection, preprocessing will be performed to enhance the quality of the data. Raw data from the industry may carry noise, missing values, and inconsistencies, which can result in false alarms or errors in detecting faults. Common preprocessing methods include signal denoising using wavelet transform,

normalization to standardize different sensor readings, and outlier removal using statistical methods. Another important step in this stage is feature extraction, which extracts relevant information from the raw data. For example, in vibration-based fault detection, time-domain features like root mean square and kurtosis and frequency-domain features like fast Fourier transform (FFT) coefficients are obtained to characterize a system's behaviour (Seid Ahmed et al., 2020).

The next phase is to choose and train the appropriate FDD model. The selection of the best model depends upon the nature of the process under consideration and the types of faults that need to be detected. ML and DL models are often used in a data-driven approach to the methodology. Classic ML algorithms such as support vector machines, K-nearest neighbors, and PCA have been around for years, used to discover deviations from the norms imposed by stealthy failures (Lu et al., 2025). Some advanced models can outperform in terms of fault classification, particularly in high-dimensional datasets, and include ANN, RNN, and CNN. Some promising results in hybrid models combine statistical signal processing and ML techniques: variational mode decomposition with deep autoencoders (Harrou et al., 2013). For model-based approaches, it is primarily concerned with providing the specifications for mathematical models that describe the dynamics of a given system. Naming extensive deviations through the help of mathematically represented systems is equations like Kalman filtering, observer-based methods, and state-space descriptions of behavior (Naya et al., 2023).

Once these models have been trained and validated, integration into industrial systems for real-time monitoring is required. This involves embedding of the trained FDD models into PLCs, distributed control systems, or cloud-based platforms. Industrial environments rely on standardized communication protocols, such as Modbus, OPC UA, and MQTT, which provide the necessary means of real-time data exchange between the FDD system and existing control systems. Denoting intelligence during the integration and deployment phase can thus be done through maximizing some measures of computational performance. Deep learning models may sometimes require computations of high value for their infrastructural needs to suffice. In that connection, some of these strategies utilize edge computing solutions, like industrial IoT (IIoT) gateways or AI-enabled microcontrollers, to perform their FDD algorithms with minimal latency.

After identifying a potential fault, the system shall move on to analysis and decision-making regarding severity, location, and impact. Various hybrid techniques are in place for rule-based decision fusion, wherein the outputs from several models are synthesized using logic-based reasoning methods. Alternatively, certain other techniques use probabilistic fusion strategies, including Bayesian inference (Bennacer et al., 2015) to moderate uncertainty in fault classification (Wu et al., 2023). Conflict resolution strategies are needed in multi-component systems for industry, when different models infer fault otherwise. One method to prevent conflicts is calculating the confidence-weighted voting, whereby a model's prediction is verified based on its measure of consistency that has been historically effective (Meyen et al., 2021). Another way is to import expert knowledge into the decision-making context to enable validation of results before actions are taken for making corrections (Hauashdh et al., 2024).

System validation, maintenance, and continuous learning are the last steps of FDD implementation. The deployed FDD system must be diligently validated with real-world industrial case studies to gauge its accuracy and reliability. FDD performance metrics, including precision, recall, and false alarm rates, can be applied for evaluating the overall effectiveness of the fault detection models. Periodic retraining of models is required to adapt to changes once an FDD model has been deployed, particularly in dynamic manufacturing environments that change over time. Adaptive online learning is used in highly advanced FDD implementations, enabling the system to self-update based on new data. Examples include reinforcement learning-based adaptive FDD that allows the model to make better predictions by incremental learning without complete retraining (Botvinick et al., 2019).

A well-implemented FDD can improve industrial reliability by detecting faults at the outset, reducing downtime, minimizing maintenance costs, and enhancing general safety in systems. However, several challenges still exist: the necessity for scalable solutions that can accommodate large volumes of industrial data, seamless integration with existing infrastructure, and the assurance of human decision-making via understandable AI-based models of FDD (Moosavi, 2024). Further advances in explainable AI, edge computing, and digital twins will significantly promote the industrial adoption of FDD methods and further demonstrate their efficiency and adaptability across various industries.

6 Directions for future research works

This paper provides a comprehensive overview of current research in the field of FDD, in addition to new developments. Fault detection and diagnosis techniques are widely used to observe processes in industrial and academic settings. However, their complexity makes it difficult to put them to practical industrial procedures. Bridging the gap between the theoretical approach and practice is necessary by accepting new hybrid methodologies and using more than one approach to create complicated FDD models (Xia et al., 2018). In further developing this domain, the following research avenues may be considered for the future.

6.1 Real-time and predictive FDD

Edge computing:

- Explore implementing low-latency FDD algorithms on edge devices for real-time decision-making.
- Investigate edge computing for data preprocessing and feature extraction before sending data to the cloud.

Prognostic health management (PHM):

- Develop advanced models for predicting components' remaining useful life (RUL) based on sensor data and historical information.
- Utilize ML techniques to improve the accuracy and reliability of RUL predictions.

Digital twins:

- Create virtual representations of physical systems to simulate and analyze fault scenarios.
- Use digital twins for model-based fault diagnosis and predictive maintenance.

6.2 Multi-sensor fusion and data fusion

Sensor fusion:

- Develop algorithms for combining data from multiple sensors to improve fault detection accuracy and robustness.
- Explore the use of sensor fusion techniques for fault localization.

Data fusion:

- Integrate data from different sources (e.g., sensors, historical data, expert knowledge) to enhance FDD capabilities.
- Utilize data fusion techniques to handle missing or noisy data.

6.3 Explainable AI for FDD

- Interpretable Models: Develop FDD models that clearly explain their decisions.
- Feature Importance Analysis: Identify the most important features for fault diagnosis.
- Model Visualization: Create visual representations of the model's decision-making process.
- Human-in-the-Loop Systems: Design FDD systems that allow human experts to provide feedback and improve model performance.

6.4 Human-in-the-Loop FDD

- Human-Machine Collaboration: Create cooperative systems involving human and artificial intelligence in detecting and diagnosing faults.
- Augmented Intelligence: Provide human operators with AI-empowered tools for improved decision-making.
- User-Centered Design: Create FDD interfaces that are intuitive and easy to operate by operators.
- Through developing these aspects, researchers and practitioners can develop more advanced and effective FDD systems that are compatible with the dynamics of manufacturing environments.

6.5 Process optimization

- Develop ML models to optimize manufacturing processes, including laser surface texturing (LST), machining, additive manufacturing, and coating applications.
- Apply ML techniques to analyze the relationship between process parameters and key performance metrics such as

surface roughness, friction, wear resistance, and material integrity.

- Use predictive modeling to enhance product quality, durability, and energy efficiency in various manufacturing applications.
- Combine ML-driven process optimization with fault detection systems to enable predictive maintenance and reduce unplanned downtime.
- Develop multi-sensor data fusion techniques to integrate process data (e.g., thermal, vibration, and tribological signals) for enhanced decision-making.
- Explore digital twins to simulate and optimize manufacturing processes before physical implementation, improving precision and reducing material waste.

7 Conclusion

The present paper has discussed the crucial roles of FDD systems in maintaining the going concern status of manufacturing systems. With a world gradually driven by Industry 4.0, an increasing need arises to apply automated FDD systems that can conduct swift and timely fault identification with subsequent suggestions for preventive measures. It systematically reviewed a total of four major FDD strategies: data-driven, model-based, knowledge-based, and hybrid-based-all offering different solutions to improve production environments in respect of efficiency, safety, and reliability. The contrasting approaches have given us directions that researchers and practitioners could use to navigate the maze of implementing FDD. Besides, hybrid-based techniques that combine strengths from different FDD strategies have great promise for overcoming limitations identified in traditional methods and hence offer strong fault detection and diagnosis capabilities.

The paper also highlights that evaluation metrics are used to prove the effectiveness of the FDD technique in real-world applications. It will be a structured approach to exploring the various methodologies of FDD, from signal acquisition and analysis to industrial applications and fault categorization, thereby providing an overview of the status and future problems. The finding has established that significant developments are realized within this field; however, further research is required for the intrinsic difficulties in applying real-time industrial settings toward FDD systems. In the future, hybrids of techniques and more sophisticated models will have to be developed to bridge the gap between theoretical advances and practical implementation for a safer, more efficient, and more productive manufacturing environment.

The FDD methods presented in a structured manner, from signal acquisition to analysis, through to industrial applications and fault categorization, provide an overview of the current state of the art on this subject and the challenges ahead. These results show that while much progress has been achieved, further research is required to overcome the innate complications of applying FDD systems to real-time industrial applications. Hybrid integration and more advanced model development over the coming years will thus be necessary to bridge theoretical advances and practical applications

toward a safer, more efficient, and productive manufacturing environment.

Author contributions

YS: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Validation, Visualization, Writing – original draft. AAA: Writing – original draft, Validation. AMA: Writing – review and editing, Methodology, Investigation. FA-B: Writing – review and editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work was supported by King Fahd University of Petroleum and Minerals, under research project EC241015. The authors would like to acknowledge the financial support provided by KFUPM for this project.

References

- Abad, M. R. A. A., Moosavian, A., and Khazaei, M. (2016). Wavelet transform and least square support vector machine for mechanical fault detection of an alternator using vibration signal. *J. Low Freq. Noise Vib. Act. Control* 35 (1), 52–63. doi:10.1177/0263092316628258
- Abaei, G., and Selamat, A. (2013). A survey on software fault detection based on different prediction approaches. *Vietnam J. Comput. Sci.* 2013 1 (2), 79–95. doi:10.1007/S40595-013-0008-Z
- Abbas, A. R., Mahmoudi, M. R., and Avazzadeh, Z. (2018). Diagnosis and clustering of power transformer winding fault types by cross-correlation and clustering analysis of FRA results. *IET Generation, Transm. and Distribution* 12 (19), 4301–4309. doi:10.1049/IET-GTD.2018.5812
- Abid, A., and Khan, M. T. (2017). Multi-sensor, multi-level data fusion and behavioral analysis based fault detection and isolation in mobile robots. *2017 8th IEEE Annu. Inf. Technol. Electron. Mob. Commun. Conf.*, 40–45. doi:10.1109/IEMCON.2017.8117139
- Abid, A., Khan, M. T., Lang, H., and de Silva, C. W. (2020). Adaptive system identification and severity index-based fault diagnosis in motors. *IEEE/ASME Trans. Mechatronics* 24 (4), 1628–1639. doi:10.1109/TMECH.2019.2917749
- Abid, A., Khan, M. T., and Silva, C. W. De (2015). 'Fault detection in mobile robots using sensor fusion 10th International Conference on Computer Science and Education, ICCSE 2015, pp. 8–13. doi:10.1109/ICCSE.2015.7250209
- Ahmed, Y. S., and Amorim, F. L. (2025). Advances in computer numerical control geometric error compensation: integrating AI and on-machine technologies for ultra-precision manufacturing. *Mach.* 2025 13 (2), 140. doi:10.3390/MACHINES13020140
- Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., et al. (2021). Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *J. Big Data* 2021 8 (1), 53–74. doi:10.1186/S40537-021-00444-8
- Ardali, N. R., Zarghami, R., and Gharebagh, R. S. (2024). Optimized data driven fault detection and diagnosis in chemical processes. *Comput. Chem. Eng.* 186, 108712. doi:10.1016/J.COMPCHEMENG.2024.108712
- Attaran, S., Attaran, M., and Celik, B. G. (2024). Digital twins and industrial internet of Things: uncovering operational intelligence in industry 4.0. *Decis. Anal. J.* 10, 100398. doi:10.1016/j.dajour.2024.100398
- Azmi, H., and Yazdizadeh, A. (2022). Robust adaptive fault detection and diagnosis observer design for a class of nonlinear systems with uncertainty and unknown time-varying internal delay. *ISA Trans.* 131, 31–42. doi:10.1016/J.ISATRA.2022.05.029
- Bakdi, A., Kouadri, A., and Bensmail, A. (2017). Fault detection and diagnosis in a cement rotary kiln using PCA with EWMA-based adaptive threshold monitoring scheme. *Control Eng. Pract.* 66, 64–75. doi:10.1016/J.CONENGPRAC.2017.06.003
- Banerjee, T. P., Das, S., Roychoudhury, J., and Abraham, A. (2010). Implementation of a new hybrid methodology for Fault signal classification using short-time fourier transform and support vector machines. *Adv. Intelligent Soft Comput.* 73 AISC, 219–225. doi:10.1007/978-3-642-13161-5_28
- Benmoussa, S., and Djeziri, M. A. (2017). Remaining useful life estimation without needing for prior knowledge of the degradation features. *IET Sci. Meas. Technol.* 11 (8), 1071–1078. doi:10.1049/IET-SMT.2017.0005
- Bennacer, L., Amirat, Y., Chibani, A., Mellouk, A., and Ciavaglia, L. (2015). Self-diagnosis technique for virtual private networks combining bayesian networks and case-based reasoning. *IEEE Trans. Automation Sci. Eng.* 12 (1), 354–366. doi:10.1109/TASE.2014.2321011
- Bernardi, E., and Adam, E. J. (2020). Observer-based fault detection and diagnosis strategy for industrial processes. *J. Frankl. Inst.* 357 (14), 10054–10081. doi:10.1016/J.FRANKLIN.2020.07.046
- Bighamian, R., Mirdamadi, H. R., and Hahn, J. O. (2014). Damage identification in collocated structural systems using structural Markov parameters. *ASME 2013 Dyn. Syst. Control Conf. DSCC 2*. doi:10.1115/DSCC2013-3758
- Bighamian, R., Mirdamadi, H. R., and Hahn, J. O. (2015). Damage identification in collocated structural systems using structural Markov parameters. *J. Dyn. Syst. Meas. Control, Trans. ASME* 137 (4). doi:10.1115/1.4028786
- Bilbao-Ubillos, J., Camino-Beldarrain, V., Intxaurburu-Clemente, G., and Velasco-Balmaseda, E. (2024). Industry 4.0, servitization, and reshoring: a systematic literature review. *Eur. Res. Manag. Bus. Econ.* 30 (1), 100234. doi:10.1016/J.IEDEEN.2023.100234
- Bordoloi, D. J., and Tiwari, R. (2014). Optimum multi-fault classification of gears with integration of evolutionary and SVM algorithms. *Mech. Mach. Theory* 73, 49–60. doi:10.1016/J.MECHMACHTHEORY.2013.10.006
- Botre, C., Mansouri, M., Nounou, M., Nounou, H., and Karim, M. N. (2016). Kernel PLS-based GLRT method for fault detection of chemical processes. *J. Loss Prev. Process Industries* 43, 212–224. doi:10.1016/J.JPLP.2016.05.023
- Botvnick, M., Ritter, S., Wang, J. X., Kurth-Nelson, Z., Blundell, C., and Hassabis, D. (2019). Reinforcement learning, fast and slow. *Trends Cognitive Sci.* 23 (5), 408–422. doi:10.1016/J.TICS.2019.02.006
- Bouchikhi, E. H. El, Choqueuse, V., and Benbouzid, M. (2015). Induction machine faults detection using stator current parametric spectral estimation. *Mech. Syst. Signal Process.* 52–53 (1), 447–464. doi:10.1016/J.YMSSP.2014.06.015
- Boudiaf, A., Moussaoui, A., Dahane, A., and Atoui, I. (2016). A comparative study of various methods of bearing faults diagnosis using the case western Reserve university data. *J. Fail. Analysis Prev.* 16 (2), 271–284. doi:10.1007/S11668-016-0080-7
- Boudinar, A. H., Benouzza, N., and Bendiabdellah, A. (2015). Induction motor cracked rotor bars fault analysis using an improved Root-MUSIC method. *International Conference on Control, Engineering and Information Technology* 1, 6. doi:10.1109/CEIT.2015.7233035
- Byun, Y. S., Kim, B. H., and Jeong, R. G. (2019). Sensor Fault detection and signal restoration in intelligent vehicles. *Sensors* 19 (15), 3306. doi:10.3390/S19153306
- Camarena-Martinez, D., Osornio-Rios, R., Romero-Troncoso, R., and Garcia-Perez, A. (2015). Fused empirical mode decomposition and MUSIC algorithms for detecting multiple combined faults in induction motors. *J. Appl. Res. Technol.* 13 (1), 160–167. doi:10.1016/S1665-6423(15)30014-6

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Chen, J., Zhang, L., Li, Y., Shi, Y., Gao, X., and Hu, Y. (2022). A review of computing-based automated fault detection and diagnosis of heating, ventilation and air conditioning systems. *Renew. Sustain. Energy Rev.* 161, 112395. doi:10.1016/j.rser.2022.112395
- Chen, Y., and Kamara, J. M. (2011). A framework for using mobile computing for information management on construction sites. *Automation Constr.* 20 (7), 776–788. doi:10.1016/j.autcon.2011.01.002
- Chen, Z. S., Rhee, S. H., and Liu, G. L. (2019). Empirical mode decomposition based on Fourier transform and band-pass filter. *Int. J. Nav. Archit. Ocean Eng.* 11 (2), 939–951. doi:10.1016/j.ijnaoe.2019.04.004
- Cheng, F., He, Q. P., and Zhao, J. (2019). A novel process monitoring approach based on variational recurrent autoencoder. *Comput. Chem. Eng.* 129, 106515. doi:10.1016/j.compchemeng.2019.106515
- Cheng, Y. J., Chen, M. H., Cheng, F. C., Lin, Y. S., and Yang, C. J. (2018). Developing a decision support system (DSS) for a dental manufacturing production line based on data mining, 638–641. doi:10.1109/ICASI.2018.8394336
- Cho, S., Gao, Z., and Moan, T. (2018). Model-based fault detection, fault isolation and fault-tolerant control of a blade pitch system in floating wind turbines. *Renew. Energy* 120, 306–321. doi:10.1016/j.renene.2017.12.102
- Choi, S. W., Lee, C., Lee, J. M., Park, J. H., and Lee, I. B. (2005). Fault detection and identification of nonlinear processes based on kernel PCA. *Chemom. Intelligent Laboratory Syst.* 75 (1), 55–67. doi:10.1016/j.chemolab.2004.05.001
- Collins, J. S., and Woodruff, G. W. (2018). 'DIGITAL TWIN VOLUME REGISTRATION FOR VOXEL-BASED CLOSED-LOOP MACHINING SYSTEMS'.
- Ding, S., Shi, Z., Tao, D., and An, B. (2016). Recent advances in support vector machines. *Neurocomputing* 211, 1–3. doi:10.1016/j.neucom.2016.06.011
- Don, M. G., and Khan, F. (2019). Dynamic process fault detection and diagnosis based on a combined approach of hidden Markov and Bayesian network model. *Chem. Eng. Sci.* 201, 82–96. doi:10.1016/j.ces.2019.01.060
- Du, X. (2019). Fault detection using bispectral features and one-class classifiers. *J. Process Control* 83, 1–10. doi:10.1016/j.procont.2019.08.007
- ElMaraghy, H., and ElMaraghy, W. (2022). Adaptive cognitive manufacturing system (ACMS) – a new paradigm', *Int. J. Prod. Res.*, 60, 7436, 7449. doi:10.1080/00207543.2022.2078248
- ElMaraghy, H., Monostori, L., I. Schuh, G., I. and ElMaraghy, W., I (2021). Evolution and future of manufacturing systems. *CIRP Ann.* 70, 635–658. doi:10.1016/j.cirp.2021.05.008
- Fan, C., Peng, Y., Shen, Y., Guo, Y., Zhao, S., Zhou, J., et al. (2024). Variable scale multilayer perceptron for helicopter transmission system vibration data abnormality beyond efficient recovery. *Eng. Appl. Artif. Intell.* 133, 108184. doi:10.1016/j.engappai.2024.108184
- Fan, J., and Wang, Y. (2014). Fault detection and diagnosis of non-linear non-Gaussian dynamic processes using kernel dynamic independent component analysis. *Inf. Sci.* 259, 369–379. doi:10.1016/j.ins.2013.06.021
- Fan, S. K. S., Hsu, C. Y., Tsai, D. M., He, F., and Cheng, C. C. (2020). Data-driven approach for Fault Detection and diagnosis in semiconductor manufacturing. *IEEE Trans. Automation Sci. Eng.* 17 (4), 1925–1936. doi:10.1109/TASE.2020.2983061
- Fathi, M., and Bolandi, H. (2024). Unsupervised optimal model bank for multiple model control systems: genetic-based automatic clustering approach. *Heliyon* 10 (4), e25986. doi:10.1016/j.heliyon.2024.E25986
- Ferentinos, K. P., Albright, L. D., and Selman, B. (2003). Neural network-based detection of mechanical, sensor and biological faults in deep-trough hydroponics. *Comput. Electron. Agric.* 40 (1–3), 65–85. doi:10.1016/S0168-1699(03)00012-7
- Galagedarage Don, M., and Khan, F. (2019). Dynamic process fault detection and diagnosis based on a combined approach of hidden Markov and Bayesian network model. *Chem. Eng. Sci.* 201, 82–96. doi:10.1016/j.ces.2019.01.060
- Gawde, S., Patil, S., Kumar, S., Kamat, P., and Kotecha, K. (2024). An explainable predictive maintenance strategy for multi-fault diagnosis of rotating machines using multi-sensor data fusion. *Decis. Anal. J.* 10, 100425. doi:10.1016/j.dajour.2024.100425
- Gnanamalar, A. J., Bhavani, R., Arulini, A. S., and Veerajay, M. S. (2023). CNN-SVM based Fault Detection, classification and location of multi-terminal VSC-hvdc system. *J. Electr. Eng. Technol.* 18 (4), 3335–3347. doi:10.1007/s42835-023-01391-5
- Golombek, R., Wrede, S., Hanheide, M., and Heckmann, M. (2011). Online data-driven fault detection for robotic systems, 3011–3016. doi:10.1109/IROS.2011.6095034
- Gonzalez, R., Huang, B., and Lau, E. (2015). Process monitoring using kernel density estimation and Bayesian networking with an industrial case study. *ISA Trans.* 58, 330–347. doi:10.1016/j.isatra.2015.04.001
- Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., et al. (2014). Generative adversarial networks. *Sci. Robotics* 3 (January), 2672–2680. doi:10.48550/arxiv.1406.2661
- Guan, B., Bao, X., Qiu, H., and Yang, D. (2024). Enhancing bearing fault diagnosis using motor current signals: a novel approach combining time shifting and CausalConvNets. *Measurement* 226, 114049. doi:10.1016/j.measurement.2023.114049
- Guo, G., Wang, H., Bell, D., Bi, Y., and Greer, K. (2003). KNN model-based approach in classification. *Lect. Notes Comput. Sci. Incl. Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinforma.* 2888, 986–996. doi:10.1007/978-3-540-39964-3_62
- Habriyansyah, I., and Andani and Zainuddin, Z. (2017). Fault detection and replacement of a temperature sensor in wheat flour short term storage, 177–181. doi:10.1109/ICA.2017.8068436
- Haddad, R. Z., and Strangas, E. G. (2016). On the accuracy of Fault Detection and separation in permanent magnet synchronous machines using MCSA/MVSA and LDA. *IEEE Trans. Energy Convers.* 31 (3), 924–934. doi:10.1109/TEC.2016.2558183
- Han, X., Zhu, H., Nie, X., Wang, G., and Zeng, X. (2018). Investigation on selective laser melting AlSi10Mg cellular lattice strut: molten pool morphology, surface roughness and dimensional accuracy. *Materials* 11 (3), 392. doi:10.3390/MA11030392
- Harandi, M. A. Z., Lin, T. Y., Li, C., Villumsen, S. L., Ghaffari, M., and Madsen, O. (2025). ScaloAdaptAlert, a novel framework for supervised anomaly detection in industrial acoustic data, integrating power scalograms, adaptive filter banks, and convolutional neural networks — a case study. *J. Manuf. Syst.* 79, 234–254. doi:10.1016/j.jmsy.2025.01.007
- Harmouche, J., Delpha, C., and Diallo, D. (2015). Incipient fault detection and diagnosis based on Kullback–Leibler divergence using principal component analysis: Part II. *Signal Process.* 109, 334–344. doi:10.1016/j.sigpro.2014.06.023
- Haroon, M., Khan, M. J., Cheema, H. M., Nasir, M. T., Safdar, M., and Butt, S. I. U. (2024). An end-to-end approach to detect railway track defects based on supervised and self-supervised learning. *Results Eng.* 24, 103326. doi:10.1016/j.rineng.2024.103326
- Harrou, F., Nounou, M. N., Nounou, H. N., and Madakyaru, M. (2013). Statistical fault detection using PCA-based GLR hypothesis testing. *J. Loss Prev. Process Industries* 26 (1), 129–139. doi:10.1016/j.jlpp.2012.10.003
- Hauashdh, A., Nagapan, S., Jailani, J., and Gamil, Y. (2024). An integrated framework for sustainable and efficient building maintenance operations aligning with climate change, SDGs, and emerging technology. *Results Eng.* 21, 101822. doi:10.1016/j.rineng.2024.101822
- He, Q. P., and Wang, J. (2007). Fault detection using the k-nearest neighbor rule for semiconductor manufacturing processes. *IEEE Trans. Semicond. Manuf.* 20 (4), 345–354. doi:10.1109/TSM.2007.907607
- Hekmat, S., and Ravanmehr, R. (2016). Real time Fault Detection and isolation: a comparative study. *Int. J. Comput. Appl.* 134 (6), 975–1012. doi:10.5120/ijca2016907931
- Heraclous, C., Miciolino, E. E., Setola, R., Pascucci, F., Eliades, D. G., Ellinas, G., et al. (2016). Critical infrastructure online fault detection: application in water supply systems. *Lect. Notes Comput. Sci. Incl. Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinforma.* 8985, 94–106. doi:10.1007/978-3-319-31664-2_11
- Hu, P., Han, Y., and Pan, J. S. (2025). An improved image-denoising technique using the whale optimization algorithm. *Electronics* 14 (1), 145. doi:10.3390/ELECTRONICS14010145
- Hu, Z., Chen, Z., Gui, W., and Jiang, B. (2014). Adaptive PCA based fault diagnosis scheme in imperial smelting process. *ISA Trans.* 53 (5), 1446–1455. doi:10.1016/j.isatra.2013.12.018
- Huang, H., Ouyang, H., and Gao, H. (2015). Blind source separation and dynamic fuzzy neural network for fault diagnosis in machines. *J. Phys. Conf. Ser.* 628 (1), 012070. doi:10.1088/1742-6596/628/1/012070
- Huang, J., and Yan, X. (2015). Dynamic process fault detection and diagnosis based on dynamic principal component analysis, dynamic independent component analysis and Bayesian inference. *Chemom. Intelligent Laboratory Syst.* 148, 115–127. doi:10.1016/j.chemolab.2015.09.010
- Jakubowski, J., Wojak-Strzelecka, N., Ribeiro, R. P., Pashami, S., Bobek, S., Gama, J., et al. (2024). Artificial intelligence approaches for predictive maintenance in the steel industry: a survey. Available online at: <https://arxiv.org/abs/2405.12785v1> (Accessed April 8, 2025).
- Jiang, B., Huang, D., Zhu, X., Yang, F., and Braatz, R. D. (2015). Canonical variate analysis-based contributions for fault identification. *J. Process Control* 26, 17–25. doi:10.1016/j.procont.2014.12.001
- Jiang, T., Gradus, J. L., and Rosellini, A. J. (2020). Supervised machine learning: a brief primer. *Behav. Ther.* 51 (5), 675–687. doi:10.1016/j.beth.2020.05.002
- Jiang, X. P., and Cao, G. Q. (2016). 'Belt conveyor roller fault audio detection based on the wavelet neural network *Proceedings - international Conference on Natural Computation*, 2016-January, pp. 954–958. doi:10.1109/ICNC.2015.7378120
- Jing, C., and Hou, J. (2015). SVM and PCA based fault classification approaches for complicated industrial process. *Neurocomputing* 167, 636–642. doi:10.1016/j.neucom.2015.03.082
- Juricek, B. C., Seborg, D. E., and Larimore, W. E. (2004). Fault detection using canonical variate analysis. *Industrial and Eng. Chem. Res.* 43 (2), 458–474. doi:10.1021/IE0301684
- Kang, S., Kwon, M., Yoon Choi, J., and Choi, S. (2023). Full-scale fire testing of battery electric vehicles. *Appl. Energy* 332, 120497. doi:10.1016/j.apenergy.2022.120497
- Kankar, P. K., Sharma, S. C., and Harsha, S. P. (2011). Fault diagnosis of ball bearings using machine learning methods. *Expert Syst. Appl.* 38 (3), 1876–1886. doi:10.1016/j.eswa.2010.07.119

- Kemalkar, A. K., and Bairagi, V. K. (2017). 'Engine fault diagnosis using sound analysis, 943–946. doi:10.1109/ICACDOT.2016.7877726
- Khireddine, M. S., Chafaa, K., Slimane, N., and Boutarfa, A. (2014). Fault diagnosis in robotic manipulators using artificial neural networks and fuzzy logic, 1, 6. doi:10.1109/WCCAIS.2014.6916571
- Kim, W., and Katipamula, S. (2018). A review of fault detection and diagnostics methods for building systems. *Sci. Technol. Built Environ.* 24 (1), 3–21. doi:10.1080/23744731.2017.1318008
- Kim, W., and Lee, J. H. (2021). Fault detection and diagnostics analysis of air conditioners using virtual sensors. *Appl. Therm. Eng.* 191, 116848. doi:10.1016/J.APPLTHERMALENG.2021.116848
- Kizza, J. M. (2020). Introduction to computer network vulnerabilities, 87–103. doi:10.1007/978-3-030-38141-7_4
- Kong, W., Luo, Y., Qin, Z., Qi, Y., and Lian, X. (2019). Comprehensive Fault diagnosis and fault-tolerant protection of in-vehicle intelligent electric power supply network. *IEEE Trans. Veh. Technol.* 68 (11), 10453–10464. doi:10.1109/TVT.2019.2921784
- Kumar, D., Ding, X., Du, W., and Cerpa, A. (2021). Building sensor Fault Detection and diagnostic system, 357, 360. doi:10.1145/3486611.3491122
- Kumar, S., Kolekar, T., Patil, S., Bongale, A., Kotecha, K., Zaguia, A., et al. (2022). A low-cost multi-sensor data acquisition system for Fault Detection in fused deposition modelling. *Sensors* 22 (2), 517. doi:10.3390/S22020517
- Kusiak, A., and Li, W. (2011). The prediction and diagnosis of wind turbine faults. *Renew. Energy* 36 (1), 16–23. doi:10.1016/J.RENENE.2010.05.014
- Kusiak, A., and Verma, A. (2012). Analyzing bearing faults in wind turbines: a data-mining approach. *Renew. Energy* 48, 110–116. doi:10.1016/J.RENENE.2012.04.020
- Kwak, J., Lee, T., and Kim, C. O. (2015). An incremental clustering-based Fault Detection algorithm for class-imbalanced process data. *IEEE Trans. Semicond. Manuf.* 28 (3), 318–328. doi:10.1109/TSM.2015.2445380
- Laouti, N., Sheibat-Othman, N., and Othman, S. (2011). Support vector machines for Fault Detection in wind turbines. *IFAC Proc. Vol.* 44 (1), 7067–7072. doi:10.3182/20110828-6-IT-1002.02560
- Lau, C. K., Ghosh, K., Hussain, M., and Che Hassan, C. (2013). Fault diagnosis of Tennessee Eastman process with multi-scale PCA and ANFIS. *Chemom. Intelligent Laboratory Syst.* 120, 1–14. doi:10.1016/J.CHEMOLAB.2012.10.005
- Li, J., Pei, H., Kochan, O., Wang, C., Kochan, R., and Ivanyshyn, A. (2024). Method for correcting error due to self-heating of resistance temperature detectors suitable for metrology in industry 4.0. *Sensors* 24 (24), 7991. doi:10.3390/S24247991
- Li, Q., Wang, Z., Wang, S., Li, M., Lei, H., and Zou, Z. (2022). A deep learning-based diagnosis model driven by tuyere images big data for iron-making blast furnaces. *steel Res. Int.* 93 (8), 2100826. doi:10.1002/SRIN.202100826
- Li, S., Peng, Y., and Bin, G. (2023). Prediction of wind turbine blades icing based on CIBM with imbalanced data. *IEEE Sensors J.* 23 (17), 19726–19736. doi:10.1109/JSEN.2023.3296086
- Lin, W. C., and Du, X. (2018). Prognosis of power connector disconnect and high resistance faults, 1, 8. doi:10.1109/ICPHM.2018.8448457
- Lin, W. C., and Ghoneim, Y. A. (2016). model-based fault diagnosis and prognosis for electric power steering systems 2016 IEEE international conference on prognostics and health management, ICPHM 2016. doi:10.1109/ICPHM.2016.7542840
- Liu, H., Wei, L., Wang, Y., Wang, J., and Li, W. (2022). Monitoring and predictive maintenance of centrifugal pumps based on smart sensors. *Sensors* 22 (6), 2106. doi:10.3390/S22062106
- Liu, J. (2012). Fault diagnosis using contribution plots without smearing effect on non-faulty variables. *J. Process Control* 22 (9), 1609–1623. doi:10.1016/J.PROCONT.2012.06.016
- Liu, Q., Chai, T., and Qin, S. J. (2012). Fault diagnosis of continuous annealing processes using a reconstruction-based method. *Control Eng. Pract.* 20 (5), 511–518. doi:10.1016/J.CONENGRAC.2012.01.005
- Liu, Z., Wang, J., Duan, L., Shi, T., and Fu, Q. (2017). Infrared image combined with CNN based fault diagnosis for rotating machinery, 137–142. doi:10.1109/SDPC.2017.35
- Lou, Z., Wang, Y., and Lu, S. (2022). A novel multivariate statistical process monitoring algorithm: orthonormal subspace analysis. *Automatica* 138, 110148. doi:10.1016/J.AUTOMATICA.2021.110148
- Lu, S., Zhou, S., Ding, Y., Kim, M. K., Yang, B., Tian, Z., et al. (2025). Exploring the comprehensive integration of artificial intelligence in optimizing HVAC system operations: a review and future outlook. *Results Eng.* 25, 103765. doi:10.1016/J.RINENG.2024.103765
- Majumder, S., and Pratihar, D. K. (2018). Multi-sensors data fusion through fuzzy clustering and predictive tools. *Expert Syst. Appl.* 107, 165–172. doi:10.1016/J.ESWA.2018.04.026
- Mansouri, M., Nounou, M., Nounou, H., and Karim, N. (2016). Kernel PCA-based GLRT for nonlinear fault detection of chemical processes. *J. Loss Prev. Process Industries* 40, 334–347. doi:10.1016/J.JLP.2016.01.011
- Mercorelli, P. (2024). Recent advances in intelligent algorithms for Fault Detection and diagnosis. *Sensors* 24 (8), 2656. doi:10.3390/S24082656
- Meyen, S., Sigg, D. M. B., Luxburg, U. v., and Franz, V. H. (2021). Group decisions based on confidence weighted majority voting. *Cognitive Res. Princ. Implic.* 6 (1), 18. doi:10.1186/S41235-021-00279-0
- Miah, M. T., Erdei-Gally, S., Dancs, A., and Fekete-Farkas, M. (2024). A systematic review of industry 4.0 technology on workforce employability and skills: driving success factors and challenges in south asia. *Economies* 12 (2), 35. doi:10.3390/ECONOMIES12020035
- Milles, A., Benbouhenni, H., Debducouche, N., Yessif, M., and Elbarbary, Z. M. S. (2024). An innovative MRAS-based technique for online detection of short-circuit faults in three-phase induction motor windings. *Electr. Eng. Prepr.* doi:10.1007/S00202-024-02826-6
- Mnassri, B., Adel, E. M. El, and Ouladsine, M. (2015). Reconstruction-based contribution approaches for improved fault diagnosis using principal component analysis. *J. Process Control* 33, 60–76. doi:10.1016/J.JPROCONT.2015.06.004
- Mostafa, S. A., Mustapha, A., Hazeem, A. A., Khaleefah, S. H., and Mohammed, M. A. (2018). An agent-based inference engine for efficient and reliable automated car failure diagnosis assistance. *IEEE Access* 6, 8322–8331. doi:10.1109/ACCESS.2018.2803051
- Mumuni, A., and Mumuni, F. (2024). Automated data processing and feature engineering for deep learning and big data applications: a survey. *J. Inf. Intell.* 3, 113–153. doi:10.1016/j.jiixd.2024.01.002
- Naqvi, M. R., Elmhadi, L., Sarkar, A., Archimede, B., and Karray, M. H. (2024). Survey on ontology-based explainable AI in manufacturing. *J. Intelligent Manuf.* 35 (8), 3605–3627. doi:10.1007/s10845-023-02304-z
- Naya, M. Á., Sanjurjo, E., Rodríguez, A. J., and Cuadrado, J. (2023). Kalman filters based on multibody models: linking simulation and real world. A comprehensive review. *Multibody Syst. Dyn.* 58 (3–4), 479–521. doi:10.1007/s11044-023-09893-w
- Nelson, W., and Dieckert, C. (2024). Machine learning-based automated Fault Detection and diagnostics in building systems. *Energies* 17 (2), 529. doi:10.3390/EN17020529
- Nguyen, H., Kieu, L., Wen, T., and Cai, C. (2018). Deep learning methods in transportation domain: a review. *IET Intell. Transp. Syst.* 12 (9), 998–1004. doi:10.1049/IET-ITS.2018.0064
- Orhan, M., and Celik, M. (2024). A literature review and future research agenda on fault detection and diagnosis studies in marine machinery systems. *Proc. Institution Mech. Eng. Part M J. Eng. Marit. Environ.* 238 (1), 3–21. doi:10.1177/1475090221149291
- Pan, L., Wu, Z., Fang, L., and Song, Y. (2019). Investigation of surface damage and roughness for nickel-based superalloy GH4169 under hard turning processing, 234(4), 679–691. doi:10.1177/0954405419885789
- Park, J., Kwon, I. H., Kim, S. S., and Baek, J. G. (2011). Spline regression based feature extraction for semiconductor process fault detection using support vector machine. *Expert Syst. Appl.* 38 (5), 5711–5718. doi:10.1016/J.ESWA.2010.10.062
- Park, Y. J., Fan, S. K. S., and Hsu, C. Y. (2020). A review on Fault Detection and process diagnostics in industrial processes. *Processes* 8 (9), 1123. doi:10.3390/PR8091123
- Pei, T., and Hao, X. (2019). A Fault Detection method for photovoltaic systems based on voltage and current observation and evaluation. *Energies (Basel)*. 12, 1712. doi:10.3390/en12091712
- Peng, B. S., Xia, H., Liu, Y. K., Yang, B., Guo, D., and Zhu, S. M. (2018). Research on intelligent fault diagnosis method for nuclear power plant based on correlation analysis and deep belief network. *Prog. Nucl. Energy* 108, 419–427. doi:10.1016/J.PNUCENE.2018.06.003
- Piltan, F., and Kim, J. M. (2018). Bearing Fault diagnosis using an extended variable structure feedback linearization observer. *Sensors* 18 (12), 4359. doi:10.3390/S18124359
- Pourbabae, B., Meskin, N., and Khorasani, K. (2016). Robust sensor fault detection and isolation of gas turbine engines subjected to time-varying parameter uncertainties. *Mech. Syst. Signal Process.* 76–77, 136–156. doi:10.1016/J.YMSSP.2016.02.023
- Qian, C., Zhu, J., Shen, Y., Jiang, Q., and Zhang, Q. (2022). Deep transfer learning in mechanical intelligent fault diagnosis: application and challenge. *Neural Process. Lett.* 54 (3), 2509–2531. doi:10.1007/S11063-021-10719-Z
- Ren, L., and Lv, W. (2014). Fault detection via sparse representation for semiconductor manufacturing processes. *IEEE Trans. Semicond. Manuf.* 27 (2), 252–259. doi:10.1109/TSM.2014.2302011
- Rubhini, B., and Ranjan, P. V. (2017). Machine condition monitoring using audio signature analysis 2017 4th International Conference on Signal Processing, Communication and Networking, ICSCN, 1, 6. doi:10.1109/ICSCN.2017.8085717
- Russell, E. L., Chiang, L. H., and Braatz, R. D. (2000). Fault detection in industrial processes using canonical variate analysis and dynamic principal component analysis. *Chemom. Intelligent Laboratory Syst.* 51 (1), 81–93. doi:10.1016/S0169-7439(00)00058-7
- Sadeghkhani, I., Hamedani Golshan, M. E., Mehriizi-Sani, A., Guerrero, J. M., and Ketabi, A. (2018). Transient monitoring function-based Fault Detection for inverter-interfaced microgrids. *IEEE Trans. Smart Grid* 9 (3), 1–2107. doi:10.1109/TSG.2016.2606519
- Saeed, A., Khan, M., Akram, U., J. Obidallah, W., Jawed, S., and Ahmad, A. (2025). Deep learning based approaches for intelligent industrial machinery health

- management and fault diagnosis in resource-constrained environments. *Sci. Rep.* 2024 15 (1), 1114–1130. doi:10.1038/s41598-024-79151-2
- Salehifar, M., Arashloo, R. S., Moreno-Equilar, J. M., Sala, V., and Romeral, L. (2014). Fault detection and Fault tolerant operation of a five phase PM motor drive using adaptive model identification approach. *IEEE J. Emerg. Sel. Top. Power Electron.* 2 (2), 212–223. doi:10.1109/JESTPE.2013.2293518
- Samy, I., Postlethwaite, I., and Gu, D. W. (2011). Survey and application of sensor fault detection and isolation schemes. *Control Eng. Pract.* 19 (7), 658–674. doi:10.1016/J.CONENGP.2011.03.002
- Sarker, I. H. (2021). Machine learning: algorithms, real-world applications and research directions. *SN Comput. Sci.* 2 (3), 160–221. doi:10.1007/s42979-021-00592-x
- Schein, J., and Bushby, S. T. (2006). A hierarchical rule-based Fault Detection and diagnostic method for HVAC systems. Available online at: <https://www.nist.gov/publications/hierarchical-rule-based-fault-detection-and-diagnostic-method-hvac-systems> (Accessed April 8, 2025).
- Sedej, O., Mbonimpa, E., Sleight, T., and Slagley, J. (2022). Artificial neural networks and gradient boosted machines used for regression to evaluate gasification processes: a review. *J. Energy Power Technol.* 4 (3), 1. doi:10.21926/jept.2203027
- Seid Ahmed, Y. (2024). Optimizing femtosecond texturing process parameters through advanced machine learning models in tribological applications. *Lubricants* 12 (12), 454. doi:10.3390/LUBRICANTS12120454
- Seid Ahmed, Y., Arif, A. F. M., and Veldhuis, S. C. (2020). Application of the wavelet transform to acoustic emission signals for built-up edge monitoring in stainless steel machining. *Meas. J. Int. Meas. Confed.* 154, 107478. doi:10.1016/j.measurement.2020.107478
- Shah, D. S., and Patel, V. N. (2014). A review of dynamic modeling and fault identifications methods for rolling element bearing. *Procedia Technol.* 14, 447–456. doi:10.1016/J.PROTCY.2014.08.057
- Shahid, M. A., Ahmad, F., Nawaz, R., Khan, S. U., Wadood, A., and Albalawi, H. (2023). A novel false measurement data detection mechanism for smart grids. *Energies* 16 (18), 6614. doi:10.3390/EN16186614
- Shang, Y., Wang, S., Tang, N., Fu, Y., and Wang, K. (2024). Research progress in fault detection of battery systems: a review. *J. Energy Storage* 98, 113079. doi:10.1016/J.EST.2024.113079
- Shao, H., Jiang, H., Zhao, H., and Wang, F. (2017). A novel deep autoencoder feature learning method for rotating machinery fault diagnosis. *Mech. Syst. Signal Process.* 95, 187–204. doi:10.1016/J.YMSSP.2017.03.034
- Su, J., and Chen, W. H. (2019). model-based fault diagnosis system verification using reachability analysis. *IEEE Trans. Syst. Man, Cybern. Syst.* 49 (4), 742–751. doi:10.1109/TSMC.2017.2710132
- Sun, X., and Jia, X. (2019). a fault diagnosis method of industrial robot rolling bearing based on data driven and random intuitive fuzzy decision. *IEEE Access* 7, 148764–148770. doi:10.1109/ACCESS.2019.2944974
- Surucu, O., Gadsden, S. A., and Yawney, J. (2023). Condition monitoring using machine learning: a review of theory, applications, and recent advances. *Expert Syst. Appl.* 221, 119738. doi:10.1016/J.ESWA.2023.119738
- Susto, G. A., Terzi, M., and Beghi, A. (2017). Anomaly detection approaches for semiconductor manufacturing. *Procedia Manuf.* 11, 2018–2024. doi:10.1016/J.PROMFG.2017.07.353
- Tabrizi, A., Garibaldi, L., Fasana, A., and Marchesiello, S. (2015). Early damage detection of roller bearings using wavelet packet decomposition, ensemble empirical mode decomposition and support vector machine. *Meccanica* 50 (3), 865–874. doi:10.1007/s11012-014-9968-z
- Tong, Z., Li, W., Jiang, F., Zhu, Z., and Zhou, G. (2018). Bearing fault diagnosis based on spectrum image sparse representation of vibration signal. *Adv. Mech. Eng.* 10 (9), 168781401879778. doi:10.1177/1687814018797788
- Tran, D. A. T., Chen, Y., Chau, M. Q., and Ning, B. (2015). A robust online fault detection and diagnosis strategy of centrifugal chiller systems for building energy efficiency. *Energy Build.* 108, 441–453. doi:10.1016/J.ENBUILD.2015.09.044
- Verdier, G., and Ferreira, A. (2011). Adaptive Mahalanobis distance and k-Nearest neighbor rule for Fault Detection in semiconductor manufacturing. *IEEE Trans. Semicond. Manuf.* 24 (1), 59–68. doi:10.1109/TSM.2010.2065531
- Vishwakarma, M., Purohit, R., Harshlata, V., and Rajput, P. (2017). Vibration analysis and condition monitoring for rotating machines: a review. *Mater. Today Proc.* 4 (2), 2659–2664. doi:10.1016/J.MATPR.2017.02.140
- Wang, G., Li, T., Zhang, G., Gui, X., and Xu, D. (2014). Position estimation error reduction using recursive-least-square adaptive filter for model-based sensorless interior permanent-magnet synchronous motor drives. *IEEE Trans. Industrial Electron.* 9 (61), 5115–5125. doi:10.1109/TIE.2013.2264791
- Wang, J., and Biljecki, F. (2022). Unsupervised machine learning in urban studies: a systematic review of applications. *Cities* 129, 103925. doi:10.1016/J.CITIES.2022.103925
- Wang, W., and Lee, H. (2013). An energy kurtosis demodulation technique for signal denoising and bearing fault detection. *Meas. Sci. Technol.* 24 (2), 025601. doi:10.1088/0957-0233/24/2/025601
- Wang, X., and Hu, B. (2024). Machine learning algorithms for improved product design user experience. *IEEE Access* 12, 112810–112821. doi:10.1109/ACCESS.2024.3442085
- Wang, Y., and Cheng, Y. (2016). An approach to fault diagnosis for gearbox based on image processing. *Shock Vib.* 2016, 1–10. doi:10.1155/2016/5898052
- Wang, Y., Ren, X., Nan, G., Yang, Y., and Deng, W. (2012). Rotating machine fault diagnosis based on denoising source separation, 1124–1127. doi:10.1109/ICACL.2012.6463348
- Widodo, A., and Yang, B. S. (2007). Support vector machine in machine condition monitoring and fault diagnosis. *Mech. Syst. Signal Process.* 21 (6), 2560–2574. doi:10.1016/J.YMSSP.2006.12.007
- Wu, H., Triebe, M. J., and Sutherland, J. W. (2023). A transformer-based approach for novel fault detection and fault classification/diagnosis in manufacturing: a rotary system application. *J. Manuf. Syst.* 67, 439–452. doi:10.1016/J.JMSY.2023.02.018
- Wu, H., and Zhao, J. (2020). Fault detection and diagnosis based on transfer learning for multimode chemical processes. *Comput. and Chem. Eng.* 135, 106731. doi:10.1016/J.COMPCHEMENG.2020.106731
- Wu, Y., Jiang, B., Lu, N., Yang, H., and Zhou, Y. (2017). Multiple incipient sensor faults diagnosis with application to high-speed railway traction devices. *ISA Trans.* 67, 183–192. doi:10.1016/J.ISATRA.2016.12.001
- Xia, M., Li, T., Xu, L., Liu, L., and de Silva, C. W. (2018). Fault diagnosis for rotating machinery using multiple sensors and convolutional neural networks. *IEEE/ASME Trans. Mechatronics* 23 (1), 101–110. doi:10.1109/TMECH.2017.2728371
- Xiao, B., Yin, S., and Gao, H. (2018). Reconfigurable tolerant control of uncertain mechanical systems with actuator faults: a sliding mode observer-based approach. *IEEE Trans. Control Syst. Technol.* 26 (4), 1249–1258. doi:10.1109/TCST.2017.2707333
- Yaman, O., Karakose, M., Akin, E., and Aydin, I. (2015). Image processing based fault detection approach for rail surface. *IEEE Trans. Syst. Man, Cybern. Syst.* 45 (1), 1118–1121. doi:10.1109/SIU.2015.7130031
- Yamashita, R., Nishio, M., Do, R. K. G., and Togashi, K. (2018). Convolutional neural networks: an overview and application in radiology. *Insights into Imaging* 9 (4), 611–629. doi:10.1007/s13244-018-0639-9
- Yan, K., Shen, W., Mulumba, T., and Afshari, A. (2014). ARX model based fault detection and diagnosis for chillers using support vector machines. *Energy Build.* 81, 287–295. doi:10.1016/J.ENBUILD.2014.05.049
- Yang, C., Li, Y., and Chen, Q. (2024). Data-driven fault detection of heterogeneous multi-agent systems using combined hardware and temporal redundant information. *ISA Trans.* 147, 90–100. doi:10.1016/J.ISATRA.2024.02.003
- Yin, Z., Hu, N., Chen, J., Yang, Y., and Shen, G. (2022). A review of fault diagnosis, prognosis and health management for aircraft electromechanical actuators. *IET Electr. Power Appl.* 16 (11), 1249–1272. doi:10.1049/ELP2.12225
- Yoshida, M., Togo, R., Ogawa, T., and Haseyama, M. (2023). Off-screen sound separation based on audio-visual pre-training using binaural audio. *Sensors* 23 (9), 4540. doi:10.3390/S23094540
- Zaben, M. M., Worku, M. Y., Hassan, M. A., and Abido, M. A. (2024). Machine learning methods for Fault diagnosis in ac microgrids: a systematic review. *IEEE Access* 12, 20260–20298. doi:10.1109/ACCESS.2024.3360330
- Zhang, C., Gao, X., Xu, T., Li, Y., Pang, Y., et al. (2019). Fault detection strategy based on weighted distance of k nearest neighbors for semiconductor manufacturing processes. *IEEE Trans. Semicond. Manuf.* 32 (1), 75–81. doi:10.1109/TSM.2018.2857818
- Zhang, L., and Zhai, J. (2019). Fault diagnosis for oil-filled transformers using voting based extreme learning machine. *Clust. Comput.* 22 (4), 8363–8370. doi:10.1007/s10586-018-1804-0
- Zhang, W. (2013). Image denoising algorithm of refuge chamber by combining wavelet transform and bilateral filtering. *Int. J. Min. Sci. Technol.* 23 (2), 221–225. doi:10.1016/J.IJMST.2013.04.016
- Zhang, Z., and Zhao, J. (2017). A deep belief network based fault diagnosis model for complex chemical processes. *Comput. and Chem. Eng.* 107, 395–407. doi:10.1016/J.COMPCHEMENG.2017.02.041
- Zhao, S., Duan, Y., Roy, N., and Zhang, B. (2024). A deep learning methodology based on adaptive multiscale CNN and enhanced highway LSTM for industrial process fault diagnosis. *Reliab. Eng. and Syst. Saf.* 249, 110208. doi:10.1016/J.RESS.2024.110208
- Zhao, Y., Ball, R., Mosesian, J., de Palma, J. F., and Lehman, B. (2015). Graph-based semi-supervised learning for fault detection and classification in solar photovoltaic arrays. *IEEE Trans. Power Electron.* 30 (5), 2848–2858. doi:10.1109/TPEL.2014.2364203
- Zheng, L., Wu, J., Zhang, S., Sun, S., Zhang, Z., Liang, S., et al. (2016). Bionic coupling of hardness gradient to surface texture for improved anti-wear properties. *J. Bionic Eng.* 13 (3), 406–415. doi:10.1016/S1672-6529(16)60313-X
- Zhou, S., Qian, S., Chang, W., Xiao, Y., and Cheng, Y. (2018). A novel bearing multi-fault diagnosis approach based on weighted permutation entropy and an improved SVM ensemble classifier. *Sensors* 18 (6), 1934. doi:10.3390/S18061934
- Zhou, Z., Wen, C., and Yang, C. (2015). Fault detection using random projections and k-nearest neighbor rule for semiconductor manufacturing processes. *IEEE Trans. Semicond. Manuf.* 28 (1), 70–79. doi:10.1109/TSM.2014.2374339

Nomenclature

FDD	Fault Detection and Diagnosis	MVA	Multivariate Analysis
FDI	Fault Detection and Isolation	CVA	Canonical Variate Analysis
FDC	Fault Detection and Classification	DWT	Discrete Wavelet Transform
FDIR	Fault Detection, Isolation, and Reconfiguration	LPV	Linear Parameter Varying
AI	Artificial Intelligence	STFT	Short-Term Fourier Transformation
ML	Machine Learning	GA	Genetic Algorithm
CPS	Cyber-Physical System	ABCA	Artificial Bee Colony Algorithm
CC	Cloud Computing	PR	Probability Ratio
DM	Data Mining	BN	Bayesian Network
IoT	Internet of Things	HMM	Hidden Markov Model
VR	Virtual Reality	NOC	Network on Chips
AR	Augmented Reality	FPGA	Field Programmable Gate Arrays
BDA	Big Data Analysis	ELM	Extreme Learning Machine
RL	Reinforcement Learning	PLS	Partial Least Squares
DL	Deep Learning	MSPCA	Multiscale Principal Component Analysis
SVM	Support Vector Machine	PM	Process Monitoring
ANN	Artificial Neural Network		
SNR	Signal-to-Noise Ratio		
FNR	Fault-to-Noise Ratio		
DVR	Data Validation and Reconciliation		
PLL	Phase Locked Loop		
BSS	Blind Source Separation		
PCA	Principal Component Analysis		
AE	Autoencoder		
RBM	Restricted Boltzmann Machines		
CV	Computer Vision		
ReLU	Rectified Linear		
RNN	Recurrent Neural Network		
GRU	Gated Recurrent Unit		
LSTM	Long Short-Term Memory		
PPO	Proximal Policy Optimization		
A2C	Advantage Actor-Critic		
NN	Neural Network		
GAN	Generative Adversarial Network		
KNN	K-nearest Neighbour		
MCMC-PF	Markov Chain Monte Carlo Particle Filter		
LSFIDI	Least Square Frequency Invariant Data Independent		
RCI	Reduction of Combined Index		
ANFIS	Adaptive Neuro-Fuzzy Inference System		
HMI	Human Machine Interaction		
ICA	Independent Component Analysis		
RBC	Reconstruction Based Contribution		