Neurocomputing 627 (2025) 129588

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Survey Paper

Data-driven machinery fault diagnosis: A comprehensive review

Dhiraj Neupane *, Mohamed Reda Bouadjenek, Richard Dazeley, Sunil Aryal

School of IT, Deakin University, WaurnPonds, Geelong, Victoria 3216, Australia

ARTICLE INFO

Communicated by W. Zhou

Keywords:
Data-driven
Deep learning
Fault detection
Federated learning
Machinery fault
Machine learning
Predictive maintenance
Reinforcement learning

ABSTRACT

In this era of advanced manufacturing, it is now more crucial than ever to diagnose machine faults as early as possible to guarantee their safe and efficient operation. With the increasing complexity of modern industrial processes, traditional machine health monitoring approaches cannot provide efficient performance. With the massive surge in industrial big data and the advancement in sensing and computational technologies, data-driven machinery fault diagnosis solutions based on machine/deep learning approaches have been used ubiquitously in manufacturing applications. Timely and accurately identifying faulty machine signals is vital in industrial applications for which many relevant solutions have been proposed and are reviewed in many earlier articles. Despite the availability of numerous solutions and reviews on machinery fault diagnosis, existing works often lack several aspects. Most of the available literature has limited applicability in a wide range of manufacturing settings due to their concentration on a particular type of equipment or method of analysis. Additionally, discussions regarding the challenges associated with implementing data-driven approaches, such as dealing with noisy data, selecting appropriate features, and adapting models to accommodate new or unforeseen faults, are often superficial or completely overlooked. Thus, this survey provides a comprehensive review of the articles using different types of machine learning approaches for the detection and diagnosis of various types of machinery faults, highlights their strengths and limitations, provides a review of the methods used for predictive analyses, comprehensively discusses the available machinery fault datasets, introduces future researchers to the possible challenges they have to encounter while using these approaches for fault diagnosis and recommends the probable solutions to mitigate those problems. The future research prospects are also pointed out for a better understanding of the field. We believe that this article will help researchers and contribute to the further development of the field.

1. Introduction

1.1. Background

Advances in science and technology, coupled with the growth of modern industry, have led to a heightened reliance on machinery, which is often operated under diverse and challenging conditions, including exposure to high humidity levels and excessive loads. These conditions can contribute to machinery failures, with significant impacts such as substantial maintenance costs, decreased production efficiency, financial losses, and, sometimes, the potential for loss of human life [1]. Therefore, the accurate and timely detection of potential machine faults and the implementation of effective maintenance strategies are essential to ensure the continued operation of machines and the safety of human lives.

Both academic and industrial communities have recognized the importance of Machinery Fault Diagnosis (MFD), leading to the development of various diagnostic methods for practical applications [2]. MFD has become an essential part of industrial development and engineering research, playing a vital role in maintaining the safety, reliability, and efficiency of critical machinery in modern industrial processes. Numerous strategies for MFD have been developed by researchers, scientists, and engineers through years of innovative and diligent work. Given the paramount importance of MFD, it is crucial to understand the different techniques of Fault Diagnosis (FD) that have been developed to address various challenges in Machine Health Monitoring (MHM). MHM is a process that involves the continuous evaluation of the overall condition of a machine to ensure optimal performance and prevent unexpected and frequent breakdowns. MFD, on the other hand, is a more focused aspect of MHM, targeting the detection and diagnosis of faults within machinery. For the purposes of this review paper, we will use the term MFD to maintain consistency and focus on the detection and diagnosis aspects.

* Corresponding author.
E-mail addresses: d.neupane@deakin.edu.au (D. Neupane), reda.bouadjenek@deakin.edu.au (M.R. Bouadjenek), richard.dazeley@deakin.edu.au (R. Dazeley), sunil.aryal@deakin.edu.au (S. Aryal).

https://doi.org/10.1016/j.neucom.2025.129588
Received 9 May 2024; Received in revised form 5 January 2025; Accepted 26 January 2025
Available online 7 February 2025

Fault diagnosis is a broad concept that includes various categories, such as fault and/or anomaly detection, fault isolation, fault identification, and fault reconstruction. Fault Detection (FDet) is the first and the easiest step in FD, focusing on identifying whether a fault has occurred. Anomaly Detection (AD) is a proactive approach that identifies unusual patterns in machinery behavior, serving as an early warning system to catch potential faults before they develop into serious issues. Fault identification involves categorizing the fault based on its characteristics, such as whether it is mechanical, electrical, or related to wear and tear. Fault reconstruction, or fault estimation, involves estimating fault amplitude using the idea of redundancy [3]. While FD focuses on identifying and addressing current faults, fault prognosis aims to predict future faults and the Remaining Useful Life (RUL) of machinery components. Prognosis methods are essential for planning maintenance activities before failures occur, thereby reducing downtime and costs [4]. RUL prediction is a key outcome of fault prognosis, providing an estimate of how long a component or system will continue to function before it fails.

To effectively utilize the insights from FD and prognosis, a systematic approach to maintenance is essential. Maintenance strategies can be broadly categorized into three types: Preventive Maintenance (PnM), Predictive Maintenance (PdM), and Reactive Maintenance (RM). [5]. PnM involves regularly scheduled inspections and part replacements to prevent failures before they occur. While effective at reducing unexpected breakdowns, PnM can sometimes result in over-maintenance, which may not be cost-efficient. RM involves taking action only after a failure has occurred. Although RM may seem less resource-intensive initially, it often leads to higher repair costs, increased downtime, and potential safety risks due to unexpected breakdowns [6]. PdM, on the other hand, uses sensor data and analytics to predict when equipment will require repair, allowing maintenance to be scheduled just in time to prevent failures. This approach minimizes downtime and extends the service life of machinery by avoiding unnecessary or premature repairs. Condition-Based Maintenance (CbM) is similar to PdM in that it relies on real-time data to determine the optimal timing for maintenance. As industrial operations move towards Industry 4.0, there is a growing trend towards PdM due to its ability to minimize costs while enhancing productivity and safety. By integrating advanced sensors and real-time data analysis, PdM, often combined with CbM, forms a robust strategy for maintaining machinery health and optimizing performance. In this review article, we focus on PdM and use the term “predictive maintenance” to refer to both PdM and CbM.

As the field of MFD continues to advance and new research emerges continuously (Fig. 1 shows the articles published in this field, source Scopus), current research trends and foundational methods are in a state of constant change. Existing reviews focusing on the application of Machine Learning (ML) and Deep Learning (DL) in MFD often deliver a fragmented summary, thus leaving significant gaps in the understanding of this field. While much of the existing literature centers on refining ML/DL models, there has been a significant shift towards understanding how fault information is learned and represented within these models. Numerous studies have surfaced offering improvements in these aspects; however, the literature lacks an overview of the specific areas and the reasons why these methods have been enhanced. Additionally, several review papers tend to provide only brief overviews of various studies, lacking in-depth exploration. These reviews often offer just one or two sentences per article, which may not offer readers a comprehensive understanding. This survey paper aims to address these issues by offering a comprehensive review of the latest research on MFD using ML and their hybrid, discussing relevant topics, and summarizing the advancements in a more detailed manner.

This paper serves as a comprehensive guide for understanding the application of various ML algorithms in the field of MFD, as well as their respective advantages and disadvantages. Additionally, it addresses the current challenges in intelligent MFD research. Intelligent

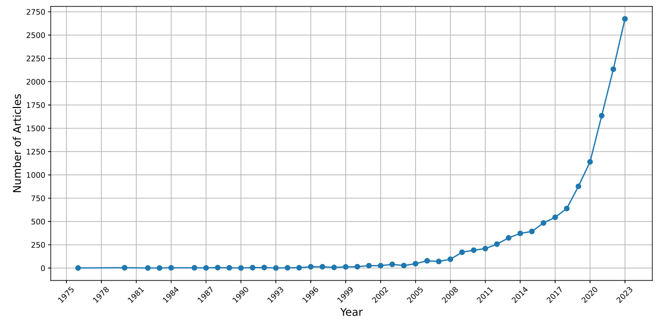

Fig. 1. Annual number of articles published in machinery fault diagnosis research (Keyword “Machinery Fault”; Source: Scopus)

MFD involves utilizing machine learning techniques to diagnose machine faults. This approach minimizes the need for human intervention and enables automatic detection of machine health states, attracting significant interest over the past two or three decades [7]. Notably, there is a scarcity of review papers dedicated to Reinforcement Learning (RL)-based approaches in MFD. Most existing research papers focus on supervised techniques and briefly touch on un/semi-supervised methods, but do not delve into RL. This gap in the literature has been thoroughly addressed in our review. One of the key distinguishing features of this review is its extensive coverage of available datasets. We provide a comprehensive overview of over 30 datasets, including table detailing relevance, usage, citations, sources, and data types. Additionally, we offer recommendations and discuss challenges for some datasets. While most reviews focus exclusively on vibration data, which remains the most commonly used data type [5], our approach is more comprehensive, encompassing a range of condition monitoring techniques, including thermal imaging, acoustic sensor data, wear debris analysis, oil analysis, and Motor Current Signature Analysis (MCSA). We have covered each technique, with some discussed in detail and others more briefly. By considering a wider range of condition monitoring techniques, we aim to bridge the gap in literature reviews and provide a comprehensive view of the field. In a nutshell, this review aims to offer a more detailed and comprehensive understanding of the current and future trends in MFD, bridge the gap in literature reviews regarding RL-based works, and provide insights into various condition monitoring techniques beyond just vibration signals.

1.2. Motivation

Given the widespread use of both DL and RL-based methods in MFD in recent years, a well-structured review of the relevant literature has become essential. Such a review would support future research in the field by providing a comprehensive summary of past studies, laying the groundwork for more in-depth explorations. There are some existing reviews summarizing the related literature from different perspectives. Zhang et al. [8], Neupane et al. [1], Hoang et al. [9], Mushtaq et al. [10], Singh et al. [11], Tang et al. [12], Sunal et al. [13], Kumar et al. [14], AlShorman et al. [15], Helbing et al. [16], Stetco et al. [17] have summarized and overviewed the ML/DL-based approaches for MFD of a particular component or equipment like rolling element bearings, gears, pumps or induction motors. Soother et al. [18] and Tang et al. [19] highlighted the importance of data processing for condition monitoring and analyzed the existing methods for data processing for DL-based MFD. Also, Zhang et al. [20] discussed the small and imbalanced data in MFD and provided the data augmentation-based, classifier design-based, and feature learning-based techniques to mitigate the issues. Similarly, Li et al. [21], Zheng et al. [22], Li et al. [2], Liu et al. [23], Hakim et al. [24] provided the review of transfer learning and domain adaptation for the machinery fault detection. Ruan et al. [25] and Pan et al. [26] focused on the Generative

Table 1
Comparison of this article with other reviews on MFD studies.

AR	Year	Machinery	Algorithms covered	Types of ML	Datasets mentioned
[16]	2018	○	U/SL	MLP, AE, DBN, CNN	○
[31]	2018	◑	U/SL	kNN, NBC, SVM, ANN, AE, DBN	○
[9]	2019	○	U/SL	CNN, AE, DBN	○
[17]	2019	○	U/SL	SVM, CNN, AE	○
[8]	2020	○	U/SL, SSL	ANN, PCA, kNN, SVM, CNN, AE, DBN, RNN, GAN, TL	◑
[1]	2020	○	U/SL, SSL, RL	AE, CNN, DBN, RNN, GAN, RL, TL	◑
[14]	2021	◑	SL	SVM, kNN, ANN, DT&RF, NBC, CNN	○
[10]	2021	○	U/SL	Mentioned TML, AE, CNN, DBN, RNN	○
[13]	2022	○	SL	SVM, MLP, RF, CNN, RNN	○
[29]	2022	○	U/SL, SSL	AE	○
[24]	2023	○	U/SL, SSL	ANN, SVM, kNN, CNN, AE, GAN, RNN, DBN, TL	◑
[32]	2023	●	U/SL, SSL	DBN, AE, RNN, CNN, GAN, TL	◑
Our Review	2024	●	U/SL, SSL, RL	TML, AE, CNN, DBN, GAN, RNN, RL, TL, FL, AD, PINN	●

Notes: ● - Fully considered ; ◑ - Partially considered; ○ - Sparsely considered. The full forms of the abbreviations used are found in Table A.13.

Adversarial Network (GAN)-based approaches for data augmentation, Anomaly Detection (AD) and fault detection and classification for machine data. Moreover, Zhu et al. [27] reviewed the applications of Recurrent Neural Network (RNN) to mechanical fault diagnosis. Yang et al. [28] and Qian et al. [29] provided a comprehensive review of Autoencoder (AE)-based approaches for industrial applications and pointed out the challenges and prospects of AE-based MFD research. Moreover, the basic structure and principles of Convolutional Neural Networks (CNNs) are discussed by Tang et al. [19] and Jiao et al. [30], focusing on analyses and summary of the applications of CNNs for fault diagnosis in rotating machinery. At the time of writing this manuscript, we could not locate any reviews specifically addressing RL for MFD. Thus, this article provides a thorough review of the latest advancements in MFD, employing Traditional Machine Learning (TML), DL, RL, Federated Learning (FL), signal-processing (SP), AD approaches, transformers, and Physics Informed Neural Networks (PINNs). Table 1 provides the comparison of our article with the other available reviews.

1.3. Organization

The structure of this paper follows the taxonomy developed for this literature review. Section 2 outlines the methodology and taxonomy created for this research. Section 3 discusses machinery fault data, including analysis and a summary of available datasets. In Section 4, we present data-driven approaches, covering both traditional and advanced methods, and highlight recent advancements in the field. This section briefly reviews conventional methods and explores various ML algorithms and learning paradigms applied in MFD. Sections 5, 6, and 7 address the challenges associated with machinery fault datasets and algorithms used for fault detection, offer recommendations, and outline future directions in the MFD field. The review concludes in Section 8. Additionally, the appendix (Appendix A, and Appendix B) includes a list of abbreviations with their full forms and a brief introduction to the extensive datasets used in this field, respectively.

2. Methodology implemented and taxonomy developed

In this review, we employed a narrative overview approach to synthesize the current state of research in MFD using ML techniques. Our primary sources of literature were Google Scholar and Scopus, where we conducted comprehensive searches using a variety of relevant keywords, including "machine fault" OR "machinery fault" OR "gear fault" OR "bearing fault" OR "induction motor fault" OR "Wind turbine fault" OR "rotor fault" OR "stator fault", etc. AND "deep learning" OR "machine learning". Additionally, we performed specific searches combining the term "machinery fault" with the names of individual ML algorithms (e.g., "machinery fault using CNN") to ensure a thorough coverage of the topic. Furthermore, we utilized the snowballing technique, which involved reviewing the references of the selected articles to identify additional relevant studies. In selecting articles for this review, we prioritized recent publications, focusing on those from venues that are recognized for publishing significant research in machinery fault detection. When encountering multiple studies on similar topics, we considered citation counts as a factor to assess the impact and relevance of the work. Moreover, for datasets search, we followed the references provided in the selected articles to locate the sources. Additionally, we conducted keyword searches (e.g., "gear fault dataset, "bearing fault dataset", etc.) to find publicly available relevant datasets for MFD.

To present a comprehensive understanding of MFD through advanced learning approaches, we constructed a taxonomy that serves as the structural foundation of this review paper. This taxonomy aims to guide readers systematically through the varied aspects of the field. Starting with an exploration of the fundamental principles of MFD and the essential background of ML, the taxonomy then directs readers through a detailed examination of diverse advanced learning techniques and their practical applications in MFD. Moreover, it extends its coverage to encompass the critical domains of data acquisition and preprocessing techniques, available datasets, and practical recommendations for creating data collection test beds. A visual representation of this taxonomy can be found in Fig. 2. In this figure, topics are categorized based on the depth of coverage in this review. Topics shaded in dark green are covered in-depth, providing comprehensive explanations and insights. Topics with partial coverage are shaded in light green, indicating a moderate level of detail. Lastly, topics that are only briefly mentioned are left unshaded.

3. Machinery fault data (and analysis)

Data is a fundamental element in all data-driven approaches, and the performance of these approaches depends significantly on the amount, quality, and diversity of the data used. Selecting the appropriate target data from the initial dataset is crucial for enhancing the reliability and accuracy of the predictive model. High-quality datasets are vital in enabling these models to learn and generalize the underlying patterns and features associated with various conditions or failures. This section discusses the importance of data for fault analysis, the ability of advanced learning algorithms to use large dataset, and strategies like data augmentation that mitigate data scarcity. Additionally, we explore the diverse data types commonly used in this field which are collected to assess and predict the condition of machinery. A comprehensive overview of the available datasets used in these analyses is provided in the Appendix B. The summary of these datasets, including their key characteristics, is presented in Table 2.

3.1. Data types and analyses

Predictive maintenance analysis for MHM relies on various types of data collected from the machine and its equipment to predict potential failures and optimize maintenance. The primary data types used in

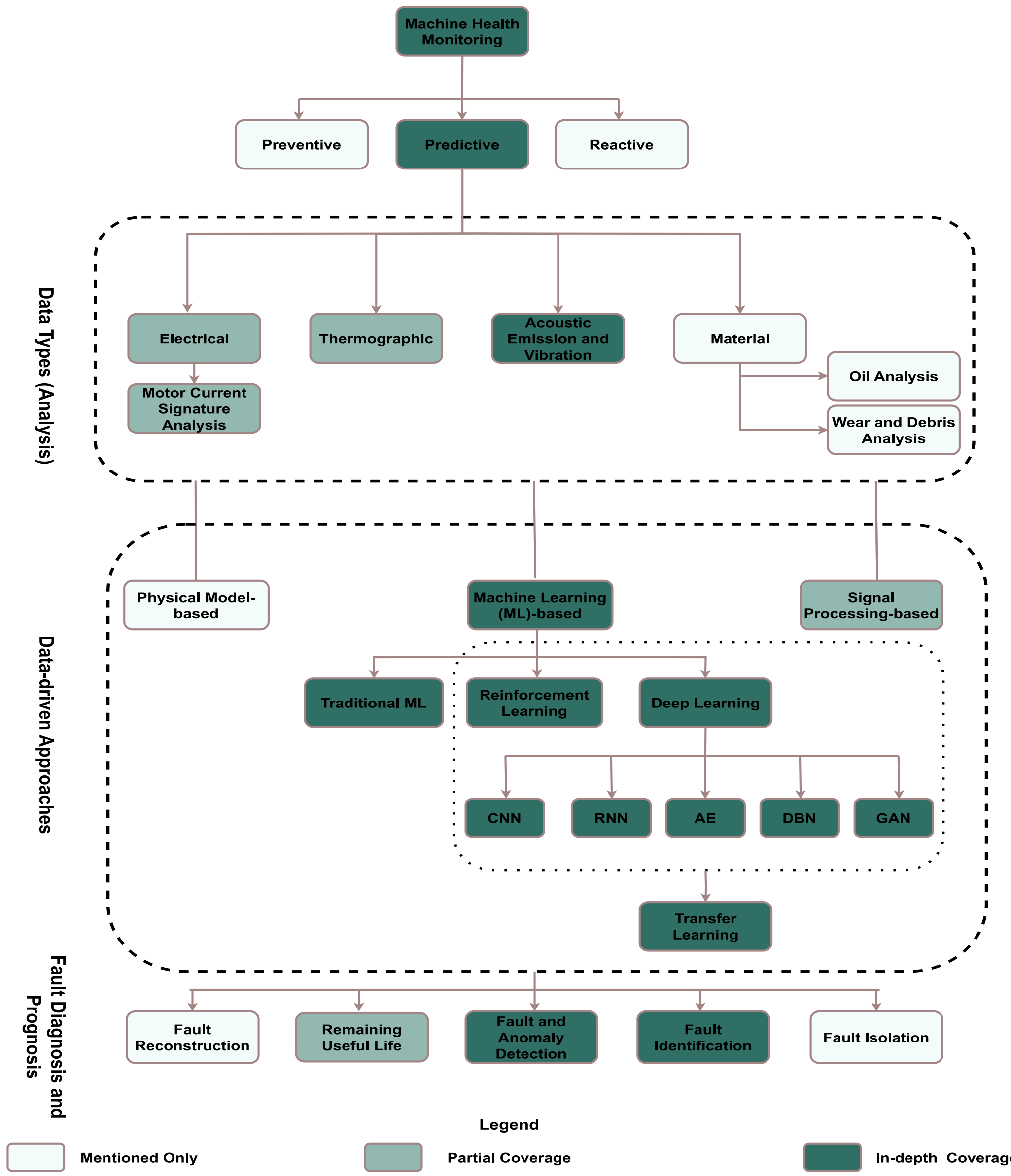

Fig. 2. Taxonomy developed for this literature review.

PdM are vibration and acoustic emission data, material data (wear & debris, oil samples), electrical data (current, phase), image data (thermal images), and other types like temperature, speed, torque, and so on. Most of these data types are time-series data, where data collection, analysis, and interpretation follow a similar pathway. Firstly, **accurate sensor placement** is critical. Sensors should be mounted close to the source to capture relevant information effectively. Additionally, secure mounting is essential to minimize interference or noise that could compromise data quality. Then, **data collection and acquisition** are facilitated by various acquisition devices–ranging from standalone to smartphone-based systems–chosen based on criteria such as the machinery's complexity, the need for portability, measurement accuracy, and budget. Ensuring that the selected device has an adequate sampling rate, resolution, and dynamic range is critical for precise data capture across all types of time-series data. The next step is **data analysis**, where the collected data is analyzed using various techniques to identify potential machinery faults. Common methods include time-domain analysis, frequency-domain analysis, and time-frequency analysis. Machine learning and deep learning techniques, as well as physics-based approaches like finite element analysis and rotor dynamics, are also employed. Statistical methods, such as calculating standard deviation, mean, kurtosis, and skewness, are used alongside data-driven approaches to enhance fault detection and diagnostics. The **ultimate goal** is to utlize these analyses to accurately detect and classify faults, localize or isolate them, and estimate the RUL of machinery. Fig. 3 represents this process.

The vibration data type is the mostly acquired and analyzed, accounting for nearly three-quarters of all PdM methods due to its effectiveness in detecting a wide range of machinery faults [33]. Other analysis methods include wear debris analysis, oil analysis, Acoustic Emission (AcE) analysis, thermography (or thermal imaging), and motor current signature analysis.

A. ***Vibration Analysis:*** Vibration analysis is a non-intrusive method sensitive to early fault stages [33]. Machines emit vibration signals that vary with condition changes, providing diagnostic information. Sensors like accelerometers capture these signals, which are then analyzed to detect and diagnose faults such as unbalance, misalignment, and bearing issues, and to predict RUL. Accurate sensor placement, data collection, and analysis are key steps in vibration analysis [162].

Table 2
Summary of datasets and corresponding articles using these datasets in MFD research.[1]

Dataset	AR	Element	Data type	Uses	Source
Airbus	[34]	Helicopter	Vibration	Classification, TFD	AirBus
AI4I	[35,36]	Synthetic	Multivariate Time-series	Classification	UCI
C-MAPSS	[37–40]	Turbofan engine	21 sensor data	Prediction, TFD	C-MAPSS
CWRU	[41–99]	Bearing	Vibration	Classification, TFD	CWRU
DIRG	[100,101]	Bearing	Vibration	Classification, TFD	Article
EDGFD	[102]	Gear	Vibration	Classification	GoogleDrive
FEMTO	[103–106]	Bearing	Temperature & Vibration	Prediction, TFD	GitHub
Gearbox Fault Diagnosis Data	[82]	Gear	Vibration	Classification	Kaggle
HUMS	[107,108]	Helicopter Gear	Vibration	Early FDet, Crack Propagation Trend, AD	HUMS2023 & HUMS2025
IMS	[43,53,69,72,99,105,109–112]	Bearing	Vibration	Classification, Prediction	IMS
IMTI	[113,114]	Induction Motor	Thermal Images	Classification	MendelyData
JNU	[115,116]	Bearing	Vibration	Classification, TFD	–
MaFaulDa	[91]	Bearing	Vibration	Classification, TFD	MaFaulDa
MFPT	[91,117]	Bearing	Vibration	Classification, TFD	MFPT
NEU	[118,119]	Steel Surface Defect	Images	Classification	thiswebsite
NREL wind Turbine	[89,120]	Gear	Vibration & Speed	Classification	OpenEI
PHM2009	[72,84,121–123]	Gearbox, bearing, shaft	Vibration	Classification, TFD	PHM2009
PHM2010	[124]	Milling Cutter	Force, Vibration, Acoustic	Prediction	PHM2010
PHME Datasets	–	Various Types	Various time series	Classification, Prediction, Root cause analysis, etc.	GitHub
PU	[49,68,82,90,101,115,117,125–127]	Bearing	Current, Torque & Vibration	Classification, TFD	PU
Rotor Fault	[128,129]	Rotor	Vibration	Classification	MendeleyData
SEU	[47,73,101,130,131]	Bearing, Gear	Vibration	Classification and TFD	SEU
TEP	[77,132–134]	Chemical process	Simulation	FDet, Process Control	TEP
THU	[41,135,136]	Gear	Vibration	Classification, TFD	–
UA-FS	[135–139]	Gear	Vibration	Classification, TFD	–
UO	[81,135]	Bearing	Vibration	Classification, TFD	MendeleyData
UoC	[71,78,79]	Gear	Vibration	Classification, TFD	UoC
UORED-VAFCLS	[140]	Bearing	Vibration, Acoustic	Classification	MendeleyData
Wind Turbine SCADA	[141–143,143–146]	Wind Turbine	Speed, Power, etc.	Classification, TFD	Zenedo, Kaggle & GitHub
XJTU-SY	[79,89,147,148]	Bearing		Prediction	XJTU-SY
MFS and Proprietary Dataset	[54–56,60–66,69,75,80,86,88,110,121,149–161]	Various elements	Various types	Classification, Prediction, TFD, etc	-

B. ***Wear Debris Analysis:*** This technique examines particles generated by machinery wear to identify faults. Analyzing the chemical makeup, size, color, and shape of wear particles identifies wear mechanisms like rubbing, cutting, and fatigue [151]. Collection methods include magnetic plugs, filters, and centrifugation, with analysis techniques like ferrography, spectroscopy, and particle counting. Effective in detecting early-stage faults in gears and bearings, wear debris analysis requires significant expertise, specialized equipment, and an understanding of machinery operation.

C. ***Oil Analysis:*** Oil analysis detects signs of wear or contamination in machinery. Regular samples are tested for properties like viscosity, acidity, particle count, and elemental composition, identifying potential faults. Common in engines, gearboxes, and hydraulic systems, oil analysis, combined with other methods like vibration analysis, provides a comprehensive understanding of machinery health [163].

D. ***Acoustic Emission Analysis:*** AcE analysis detects high-frequency sound waves generated by defects or wear. AcE sensors attached to machinery detect signals processed to extract information. Common in rotating machinery and structures, AcE analysis identifies defects through sound wave patterns [164].

E. ***Thermography:*** Thermography uses infrared cameras to capture temperature distribution, indicating potential faults. Infrared cameras detect emitted radiation, creating images representing temperature variations. Analyzing these images identifies hot spots or anomalies in electrical systems, motors, gearboxes, bearings, and insulation materials [165].

F. ***Motor Current Signature Analysis:*** MCSA monitors electric motors by analyzing the current signal in motor windings, detecting faults like bearing, rotor, and stator issues. The signal spectrum analysis identifies operating condition anomalies, helping diagnose faults and optimize maintenance [166].

3.2. Data augmentation

Acquiring comprehensive datasets for MFD is a challenging task due to the rarity of failures and high data collection costs, which often result in limited and biased datasets [167]. To address this issue, various techniques, such as data augmentation and transfer learning, are employed to enhance dataset diversity and improve model performance. Data augmentation involves transforming existing data in the time or frequency domains to generate new samples, effectively preventing overfitting and improving generalization [168]. On the other hand, Transfer Learning (TL), which is explained in Section 4.4, uses related

[1] The full forms of these abbreviations can be found in Appendix A.

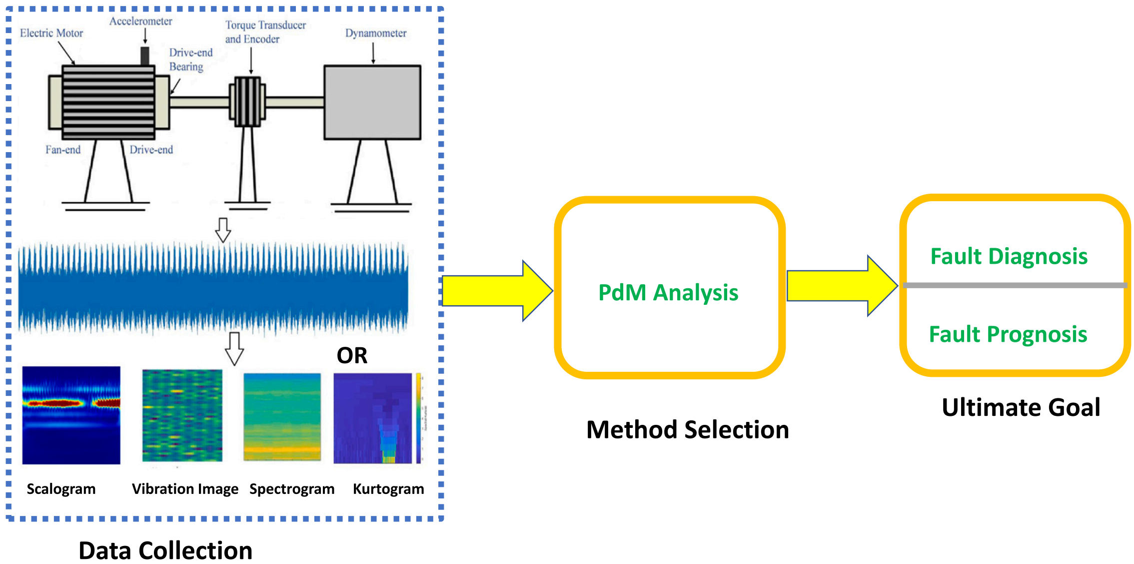

Fig. 3. General framework of MFD using ML-based algorithms.

Table 3
Data augmentation Techniques Used in MFD.

Augmentation techniques	Article references
Oversampling	SMOTE [169], SCOTE [170], K-means SMOTE [171], SI-SMOTE [172], imputation-based [173], cluster-majority weighted minority-based [174], Overlapping [175], clipping [176], Gaussian Noise [19,177]
Data Transformation	Cut-flip [178], time warping [179,180], permutation [180]
Generative	cGAN [181], GFMGAN [182], WCGAN [183], DCGAN [184], Wasserstein GAN [185], VAE [186], cVAE-GAN [187], VMR [188], cVAE [189]

domain knowledge. Moreover, Collaboration and data sharing among researchers are also encouraged to promote the development of robust MFD models [48]. These approaches help to overcome the issues of data scarcity and imbalance that can affect the accuracy of data-driven fault classification methods. A detailed list of studies that have employed these techniques can be found in the table Table 3.

4. Data-driven approaches

In MFD, the PdM techniques can be broadly classified into Physical Model (PM)-based [190], Expert System (ES)-based, Signal-Processing-based [191], ML-based [7], and their hybrids [23], according to their developmental progression. However, for the purpose of this literature review, we have organized them into two distinct categories: (i) Traditional Data-driven Approaches, encompassing PM-based, ES-based, SP-based, and TML-based, and (ii) Advanced Data-driven methods, which include DL, RL, TL & Domain Adaptation (DA) and hybrid approaches.

4.1. Traditional data-driven approaches

A. **Physical Model-based Approach:** PM-based methods rely on the system's mathematical or analytical model. It is like knowing the blueprint of the machine. The task involves identifying defects in the processes, actuators, and sensors by leveraging the dependencies among various measurable signals [192]. PM-based methods are inefficient and inflexible for practical use because (i) they require a thorough understanding of machine mechanism, (ii) it is challenging to build accurate physical systems for modern complex mechanical devices operating in noisy environments, and (iii) they are unable to update data in real-time.

B. **Expert System-based Approach:** Similarly, MFD through ES-based approaches relies on the automation of the diagnostic process using expert-level knowledge. These systems employ various reasoning methods, including rule-based reasoning, fuzzy logic-based reasoning, neural network-based reasoning, and case-based reasoning [7]. Despite their advantages, these methods face limitations, such as their heavy reliance on expert knowledge, which is difficult to acquire and express, and their lack of self-learning capabilities, which restricts their ability to expand and update the knowledge base.

C. **Signal-processing-based Approach:** Furthermore, the SP-based methods aim to extract the relevant information from the collected signals to identify the potential faults in a machine. These techniques help emphasize the machine's fault status characteristics by employing advanced signal filtering and denoising methods [193]. The commonly used SP methods in MFD are Fast Fourier Transforms (FFT) [194], Wavelet Transforms (WT) [195], Wavelet Packet Transform (WPT) [80] Empirical Mode Decomposition (EMD) [196], Hilbert-Huang Transform (HHT) [197], cepstrum analysis [198], envelope analysis [199], variational mode decomposition [200], and so on. With the use of these techniques, the appropriate level of accuracy can be achieved; however, (i) these methods struggle with complex, non-linear, or non-stationary signal sources, (ii) are time-consuming, (iii) have limited predictive capability for prospective faults, and (iv) usually require a mathematical basis and deep technical knowledge for feature extraction and understanding the significance of different frequency components.

D. **Traditional Machine Learning-based Approach:** TML algorithms have been widely used in MFD. Initially, in the early 2000s, these algorithms were primarily employed for fault-state classification or identification processes. Over time, their application expanded to include integration with the SP methods and, eventually, their incorporation into deeper networks as ensemble networks. The commonly used TML algorithms for MFD are Artificial Neural Network (ANN), Support Vector Machine (SVM), Decision Tree and Random Forest (DT & RF), K-Nearest Neighbor (KNN), Naive Bayes Classifier (NBC), Hidden Markov Model (HMM), K-means clustering, C-means clustering, Principal Component Analysis [201], regression analysis [202], and so on. Table 4 presents some of the works employing TML in the field of MFD and also presents the general advantages and disadvantages of these algorithms.

4.2. Deep learning-based data-driven approaches

While TML-based techniques show promise in MFD, they struggle with early fault detection in complex industrial environments. The dynamic, non-linear, and multi-modal nature of industrial processes complicates data analysis. TML methods require manual feature extraction and selection, which is challenging for large datasets and separates feature mining from decision-making, leading to inefficiencies [250]. As machine complexity and data dimensions grow, these limitations hamper the effectiveness of classical models. Additionally, traditional fault detection methods, dependent on human expertise, are time-consuming, error-prone, and inadequate for modern industrial systems [27]. The presence of noise and non-stationary machinery signals further complicates fault detection. Consequently, DL has gained significant interest in improving fault detection and diagnosis. Emerging as a powerful approach in the mid-2010s [251], DL has revolutionized MFD by learning useful representations directly from the raw data, eliminating the need for manual feature engineering. This capability, combined with advancements in sensors and IoT, allows DL to process large volumes of data, improving the accuracy and efficiency of fault detection and diagnosis [29,252].

FD in rotating machinery using advanced data-driven approaches builds on the steps outlined in Section 3.1 (also represented in Fig. 3) The steps include: (a) collecting sensor data that can indicate the health of the equipment, (b) extracting features from the collected data using various models and algorithms, and, (c) identifying and classifying the fault state of the equipment based on the extracted fault-sensing features [32]. In these approaches, vibration, acoustic emission, temperature, and current data are commonly used, with specific methods applied to extract features from different domains (time, frequency, time-frequency). Feature selection methods then identify the most relevant features, which are used to train ML models for fault classification, anomaly detection, and RUL prediction. This section explores the principles, techniques, and applications of different types of DL algorithms used in MFD.

4.2.1. Convolutional neural network-based FD

CNNs are widely used for machine fault classification. In fact, MFD is one of the earliest and most extensively explored fields in CNN-based FD (CNNFD), directly inspired by the image classification principle [30]. Both 2-D and 1-D CNNs are applied in this context. 1-D CNNs focus on time series data, while 2-D CNNs handle multidimensional data with spatial/temporal correlations. Different variants of CNNs, like ResNet, DenseNet, VGGNet, capsule networks (CapsNet), dilated CNN, region-based CNN, and more, have also been adapted and deployed in MFD to enhance manufacturing and industrial processes. For an organized review, we have subdivided the study of CNNFD based on the structural characteristics of convolutional networks into two aspects: FD using 1-D CNN and 2-D CNN.

A. **1-D CNNFD:** Employing 1-D CNN for FD is a straightforward strategy in which the raw 1-D data can serve as a direct input to the CNN model. Vibration data, as previously mentioned, is the prevalent data type within this strategy. The study [111] employed a 1-D CNN (3 convolution and 2 multi-layer perceptron layers) for bearing fault classification of the IMS dataset. The raw vibration data undergoes several essential preprocessing steps. It is decimated by a factor of 8 to manage complexity, providing a bandwidth of 12.5 KHz. Low-pass filtering eliminates high-frequency components, and data is normalized for consistent scaling. The study achieved 97.1% of accuracy. Similarly, Zhang et al. in [76] employed 1-D CNN for bearing fault diagnosis using CWRU bearing dataset under a noisy and variable working environment. Utilization of dropout and small-batch training techniques, and avoidance of complex prepossessing makes this work unique. Slicing the training samples with overlapping techniques is utilized as a data augmentation method. An impressive classification accuracy of 99.77% was achieved, particularly under a signal-to-noise ratio of 10. Additionally, the t-SNE method

Table 4
Summary of applications of TML-based algorithms used in MFD, and their respective advantages and disadvantages.[2]

AR	M	Advantages	Disadvantages	Model
[203–207]	B	High self learning capability	Increased complexity with more data	
[208–210]	G	Recognizes multiple machine states	Potential overfitting	ANN
[211–214]	IM		Low interpretability	
[215,215,216]	B	Better interpretability compared to ANNs	Sensitive to kernel parameters	
[217–219]	G	High diagnosis accuracy	Complicated for multi-class tasks	SVM
[220,221]	IM		Inefficient with massive data	
[222,223]	B	Easy implementation	High computation cost for large datasets	
[217,224]	G	Intuitive and simple to understand	Difficult parameter (k) determination	KNN
[220,225]	IM		Reduced accuracy with imbalanced data	
[226–228]	B	Handles missing data	Prone to overfitting	
[229,230]	G	Naturally interpretable	Often constructed based on expert knowledge	DT & RF
[231,232]	IM			
[233,234]	B	Simple, fast and scalable	Assumes feature independence	
[229,235]	G	Low training complexity	Poor performance with highly correlated features	NBC
[231,236]	IM			
[237,238]	B	Captures temporal dependencies	Computationally intensive	
[239,240]	G	Can handle noisy data	Sensitive to initial parameters	HMM
[241,242]	IM			
[243–245]	B	Summarizes data and reduces dimensions	May not scale well with large datasets	
[246,247]	G	Understands natural groupings	Some algorithms are computationally expensive	Clustering Methods
[248,249]	IM			

[b] AR: Article references; B: Bearing; G: Gear; IM: Induction Motor; M: Machinery type used.

was employed to visualize and comprehend the classification results. Moreover, in the study [77], a zero-shot FD method based on semantic space embedding was implemented. This approach utilized a 1-D CNN with two convolution layers and a fully connected layer for extracting fault features from raw data from CWRU and TEP datasets. To enable zero-shot learning, human-defined fault label embeddings were created as a fault attribute matrix with seven fine-grained fault attributes for each bearing fault. Feature embeddings were matched with fault attributes using cosine distance. Moreover, the study [112] utilized the IMS bearing vibration dataset to enhance electric motor bearing fault detection via a multichannel 1-D CNN classifier processing time-domain vibration data. Z-score standardization and linear scaling were applied to preprocess the data. The two-channel classifier achieved 100% accuracy by using both x and y-axis data, and a two-channel/two-level classifier distinguished early and advanced fault levels with 84.64% average accuracy. Similarly, in [82], a Wide-Kernel CNN (WK-CNN) was utilized to process raw time series data from three industrial machinery fault datasets: CWRU bearing, gearbox fault diagnosis, and Paderborn bearing datasets. The methodology focused on varying architectural hyperparameters (like kernel size, stride, and filters) across 38,880 model iterations, analyzing their impact on model performance. Data preprocessing involved segmenting, batch processing, and splitting into training and test sets. The results, measured primarily through accuracy scores, revealed significant performance variations across datasets and hyperparameter configurations, with some models achieving near-perfect accuracy. Additionally, ML/DL algorithms were employed to further understand the non-linear relationship between specific hyperparameters and the WK-CNN's performance, underscoring the nuanced influence of the number of trainable parameters on different datasets.

Apart from vibration data, other mechanical fault signals, like AcE, current, temperature and so on, are also used as an input for 1-D CNNFD. In [253], authors presented a blade damage identification method for centrifugal fans. They used a multi-level fusion algorithm, combining vibroacoustic signals–collected from acoustic pressure sensors and accelerometers under different fan speeds and noise levels–with a 1-D CNN network. Acoustic and vibration signals were fused at the data level using adaptive weighted fusion, followed by feature extraction with a 1-D CNN network. The extracted features were then combined through a fully connected layer. Similarly, the work [141] focused on predicting wind turbine blade icing faults by combining ReliefF feature extraction with a 1-D CNN stacked bi-directional gated recurrent unit model. Using SCADA data from China's 2017 industrial big data competition, initial data preprocessing involved reducing feature dimensions from 20 to 15 using ReliefF and reconstructing the data with a sliding time window. A weighted accuracy metric addressed data imbalance, and 5-fold cross-validation was used. Results showed a significant 43.08% increase in weighted accuracy compared to traditional models. The study [254] aimed to diagnose faults in permanent magnet synchronous motors by analyzing stator current signals. Stator current data from various fault scenarios and speeds were collected, resulting in a 2000-point dataset, which was split into 1000 for training and 1000 for testing after normalization. Two key methods, the 1-D CNN and the WPT, were employed for feature extraction. The 1-D CNN outperformed other methods with a diagnostic accuracy of 98.8%.

With the increasing concept of Explainable Artificial Intelligence (XAI), the need for transparent and understandable models in complex mechanical systems is more critical than ever. Addressing this, the study [101] utilized a Multi-Wavelet Kernel (MWK) CCNN to analyze 1-D vibration signals from gearboxes. It employed DDS datasets provided by Southeaset University, China and Wind Turbine Gearbox Dataset. Integrating a CWT with a traditional CNN, this study introduced a MWK convolutoin layer and a kernel weight recalibration module, and employed heatmaps for visualizing learned feature maps, thereby improving the interpretability of impulse detection in gearbox fault diagnosis.. Preprocessing of the data involved signal cutting using a sliding window of size 1024 and standard normalization techniques. The MWK-CNN achieved over 98% classification accuracy in gearbox fault diagnosis under various conditions.

Table 5
2-D CNN applications in MFD according to types of inputs.

Transformation Methods	Article references
Data Matrix	[237,257–266]
Vibration Image	[42,267–274]
Thermal Image	[255,275–278]
Time/Frequency Transformation	[83,167,256,279–285]
Short-time Fourier Transform	[286–291]
Wavelet Transform	[45,292–301]

B. **2-D CNNFD:** Initially, CNN architectures applied for MFD imitated the 2-D structure used in image processing. Since mechanical signals are mainly 1-D time series, the main approach was to change these 1-D signals into 2-D. To do this, several SP methods were implemented, which are described briefly in the following paragraph.

Many researchers convert 1-D time series data into 2-D image formats using data matrix transformations, referred to as "vibration images" [42]. Here, the raw 1-D data is transformed to 2-D matrices (image-structure). For example, if we have a 1-D signal $X = [x_1, x_2, \ldots, x_m]$ and wish to convert it into a 2-D matrix of dimensions $a \times b$ (where $a \times b = m$), the transformation can be applied as $Y_{i,j} = x_{(i-1)\cdot b+j}$ for $1 \leq i \leq a$ and $1 \leq j \leq b$. This process effectively restructures the time series data into a format similar to an image, allowing for the application of image processing techniques to analyze the signal patterns. Besides "vibration images", thermal images [255] have also been used as input for CNNs. In MFD, thermal imaging is key for spotting temperature anomalies that indicate faults, enabling non-invasive, continuous monitoring and diagnosis, often used alongside other techniques like vibration analysis to prevent major machine issues. Additionally, kurtograms [256], Scalograms [45], and Spectrograms [42] are important SP techniques in MFD, which analyze machine signal and noise. Kurtograms focus on detecting transient faults by assessing signal kurtosis across frequency bands, highlighting sudden signal changes. Scalograms derived from the continuous wavelet transform (CWT) provide a time-frequency view of the signal that efficiently highlights frequency changes over time. Spectrograms generated using the short-time Fourier transform (STFT) provide a time-frequency analysis ideal for tracking changes in machine behavior. These techniques facilitate the detection of machine faults and potential defects by processing raw vibration or acoustic data through mathematical algorithms and converting the signal into a visual form. The data types used in 2-D CNN are represented in Table 5.

As an application of data matrix transformation method, Neupane et al. [42] used this technique to transform raw data to 2-D matrices and then employed a simple 2-D CNN model for the bearing fault detection of CWRU dataset. Similarly, in their innovative study, Jiao et al. [302] proposed a deep coupled dense convolutional network (CDCN) for mechanical fault diagnosis. Using a 1-D convolutional structure and dense connections, the CDCN effectively extracts features from raw, nonstationary mechanical signals. Unlike conventional data-splicing techniques, the model integrates multisensor data as parallel inputs. These sensors capture both transverse and torsional vibrations, enabling a double-level information fusion approach for more accurate fault identification. The CDCN model, achieving 99.39% accuracy, demonstrated superior recognition accuracy, convergence speed, and classification accuracy compared to traditional CNNs, single-sensor data methods, and data-splicing-based fusion techniques. All the experiments were carried out with an Intel Core i7 CPU and GEFORCE GTX 1060 GPU 10 times to reduce randomness. The authors simulated and tested nine health conditions of a planetary gearbox on a test rig, running experiments at 20 Hz driving speeds with a 2 N⋅m load, and collected data at a sampling frequency of 5 kHz, resulting in a total dataset of 5400 samples for the nine conditions; they then add Gaussian white noise to simulate a real, harsh industrial environment. Even though the model could potentially increase computational costs and introduce discrepancies between training and testing data distributions, its promising results and planned improvements indicate that this approach has contributed substantially to the progress in the field of intelligent fault diagnosis.

Alternatively, employing the statistical data from either the time or frequency domain as the input for a convolutional network is another method of conversion. In the study [303], a gear test rig with an accelerometer and high-speed camera was used to train a CNN (VGG16 ConvNet) with 2-D grayscale images from FFT spectrums of vibration signals. The dataset included 600 images, 500 for training and 100 for testing, from endurance tests on plastic gears. Employing transfer learning, the VGG16, initially trained on ImageNet, was retrained with these vibration data images. In the preprocessing step, vibration signals were transformed into images using an FFT spectrum peak picking method. This method involved selecting amplitude peaks across frequencies from 0 Hz to 1600 Hz in a zig-zag pattern, every 16.67 Hz, resulting in 12×16 pixel grayscale images. Each image, representing frequency amplitudes and phases, was then labeled as ‘crack’ or ‘non-crack’ based on high-speed camera observations, providing labeled data for CNN training. Remarkably, the study achieved 99% accuracy in training and 100% in testing, primarily aimed at detecting cracks in plastic gears, with a focus on retraining the model’s final two layers. Moreover, the research detailed in [83] introduces a FD method for rolling bearings, integrating a deep CNN (DCNN) with an immunity algorithm. This technique uses 2-D images derived from bearing fault signals in both time and frequency domains as input for the DCNN. Utilizing data from the CWRU Data Center, the method combines DCNN for feature extraction and an immunity algorithm for adaptive learning of new faults. Preprocessing includes transforming vibration signals into 2-D images. The method achieves over 98% accuracy in fault recognition, effectively minimizing false positives and negatives, showcasing its proficiency in adaptive learning and accuracy in fault diagnosis through advanced machine learning. Moreover, the study [256] focuses on FD in rolling-element bearings using CNN, specifically a modified LeNet-5 architecture. It uniquely transforms 1-D AcE signals into 2-D kurtogram images for CNN compatibility. The dataset comprises AcE signals from bearings with various fault conditions and normal states. The preprocessing includes converting 1-D signals into 2-D kurtograms using a 1/3-binary tree approach. The 1/3-binary tree approach in the fast kurtogram algorithm is an extension that uses three additional band-pass filters to further decompose signal sequences, achieving finer frequency and resolution sampling with negligible extra computing cost. With this methodology, classification accuracies for different bearing conditions at various speeds exceeded 95%, with some conditions achieving 100% accuracy at 500 RPM. Furthermore, the study [81] introduces AntisymNet, a lightweight CNN for diagnosing faults in rotating machinery, transforming 1-D vibration signals into 2-D images for processing. Utilizing datasets like MiniImageNet, CWRU Bearing, Ottawa Bearing, and Hob datasets, the model achieves high accuracy (up to 99.70% on CWRU). AntisymNet’s innovative architecture combines forward and reverse branches for efficient feature extraction and fusion, demonstrating reduced complexity and strong performance across different data ratios, underscoring its practical application in industrial fault diagnosis. Also, in the study [304], a 0.5 HP induction motor connected to a SpectraQuest MFS was analyzed using two accelerometers. The study employed a multi-head 1-D CNN, which analyzed 1-D vibration signals to diagnose motor faults, including bent shaft and bearing issues. The data, divided into 256-sample windows, was collected from four experimental runs. The CNN, equipped with Leaky ReLU and Early Stop features, achieved a 99.92% fault recognition accuracy.

Apart from fault detection, some research like [105] focused on the RUL prediction of rolling element bearing using CNN architecture based on similarity feature fusion. Using FEMTO and IMS dataset for validating the model, the methodology involved feature extraction, construction of similarity features based on Pearson correlation coefficient, selection of high-sensitivity features, construction of health index using PCA and RUL prediction using 1-D CNN. Preprocessing included wavelet denoising, moving average filtering, normalization, and outlier elimination.

Apart from these image types used, some researchers have employed thermal images as the input to the CNN. Li et al. [255] developed a fault diagnosis method using CNN for infrared thermal (IRT) images, captured from the machinery applying the IRT technique. The IRT images, acquired from the SpectraQuest machinery fault simulator, are selected from the thermal video to construct the data samples, and they are fed into the CNN model for fault detection. The number of fault classes used is 10, and SoftMax is used as the classifier. Similarly, in their work [165] on machine fault diagnosis, researchers employed a robust bearing test rig with a 220 V, 2 HP DC motor to simulate various bearing faults like inner race, outer race, ball faults, and lubrication issues. Data was collected using a uniaxial accelerometer for vibration and a thermal imaging camera for IRT. They adopted 2-D CNN, processing 1-D vibration data into 2-D images through CWT for scalograms and extracting thermal images via IRT. The data, converted to grayscale for computational efficiency, enabled the CNN to achieve 100% accuracy in fault diagnosis under constant speed conditions, with slightly reduced accuracy during speed variations.

Employing 2-D CNN with XAI principles, the study [101] introduces a multilayer wavelet attention CNN (MWA-CNN) for MFD, merging CNNs with wavelet transform techniques. It employs a discrete wavelet transform layer and frequency attention mechanism to enhance noise robustness and interpretability. The methodology alternates between DWT and convolutional layers for signal decomposition and feature learning. Data from high-speed aeronautical and motor bearings (DIRG and PU bearing datasets) were used, with Z-score normalization for preprocessing and sliding segmentation for data augmentation. Notably, MWA-CNN achieved high diagnostic accuracies (98.75% at 4 dB SNR and 87.61% at −4 dB SNR). This approach improves interpretability, aligning with XAI principles, by enabling the network to focus on relevant feature information, thus enhancing decision-making transparency. Some of the other works using CNN are summarized in Table 6.

4.2.2. Recurrent neural network-based FD

RNNs are designed to handle time-series data by processing inputs sequentially and using feedback loops to remember past states, making them ideal for analyzing temporal characteristics of machinery signals like vibration and temperature. This ability enables effective fault detection and prediction of RUL. However, RNNs often struggle with vanishing and exploding gradients, which limits their capacity to learn from long sequences [1]. This has led to the development of advanced versions such as Long Short-Term Memory (LSTM) units and Gated Recurrent Units (GRUs), which incorporate gated mechanisms

Table 6
Articles employing CNN in MFD.[3]

AR	M	Dataset	Algorithm(s)	Result	Remarks
[160]	WTG	Proprietary	Wavelet Packet Decomposition combined with a hierarchical CNN	testing accuracy of approximately 98.45% and inference time in milliseconds	Dataset:14,240 segments, 2048 data points each, 5120 Hz sampling, diverse gearbox health conditions
[258]	B	CWRU	Adaptive DCNN (2-D)	92.84% without PP and 99.71% with PP	Decomposed vibration signals into 8 frequency bands, selected most sensitive for input; PSO method used for determining the main parameters of CNN
[85]	B, G	CWRU, UESTC	Dual convolution capsule network (1-D)	1̃00% accuracy with CWRU dataset	Wide convolution kernel in first layer and narrow kernels in second; Employed k-fold cross-validation in training
[167]	B	Proprietary	CNNs (2-D)	93.16% with CNN and 87.25% with RF and SVM	Scaling accelerometer signals, windowing the data, and applying the Discrete Fourier Transform (DFT) to the signals
[122]	G	PHM2009	CNN embedded with a health-adaptive time-scale representation	99.24%, using the hyperparameters of 64 scales and 8 channels	Developed HTSR method to convert 1-D vibration signals into 2-D format; Employed multiscale 1-D convolutional filters as adaptive basis functions
[305]	G	Simulated helical gearbox fault dataset	Compressive sensing-based dual channel CNN (2-D)	Average accuracy of 99.39%	Utilized Modulation Signal Bispectrum for acoustic signal analysis and converted thermal images to grayscale, computing MSB magnitudes from these signals; Utilized a mobile phone for non-contact measurements of thermal images and acoustic signals
[257]	B	CWRU	Hierarchical Adaptive DCNN (2-D)	99.7% testing accuracy	1-D raw signal to 2-D image matrices manipulation
[268]	B	CWRU	CNN	100% (with 30 and 60 kernel size in the first and second convolutional layer, respectively)	vibration images are created using matrix manipulation; Evaluation with varying kernels in convolution layers, different load conditions, and noise conditions
[86]	B	CWRU, Proprietary (AT)	weight-shared capsule network (WSCN), optimized with a margin loss function and an agreement-based dynamic routing algorithm (1-D CNN)	98.6% accuracy (CWRU)	WSCN: longer training time than traditional models; Weight-shared architecture: Fewer parameters, quicker training, better efficiency
[87]	CP, B	Proprietary, CWRU	Hierarchical symbolic analysis combined with CNN (1-D)	98.50% accuracy in a multi-load dataset	PP: Data decomposed into frequency components through hierarchical analysis; Comparative analysis includes STFT-CNN and CWT-CNN methods
[306]	WTG	Simulation	Multiscale CNN (1-D)	96.76% to 99.59% for different scales	Coarse-graining of raw vibration signals to represent them at multiple time scales
[88]	B, R	CWRU, Proprietary	Hierarchical training using 2-D CNNs with magnet-loss pretraining for parameter initialization and feature distribution, followed by global fine-tuning	96.56% (CWRU), 94.28% (Proprietary)	PP: Noise filtering, abnormal data exclusion, frequency domain conversion, and computation of 15 statistical features from vibration data

to improve learning from long-term dependencies, crucial for reliable fault diagnosis in machinery.

- **LSTMs and GRUs:** LSTMs are designed to solve the vanishing gradient problem of RNN. They employ a more complex cell structure with a memory cell, input gate, output gate, and forget gate, allowing them to learn and retain long-term dependencies in the input sequences. GRUs (Gated Recurrent Units) use a simpler architecture than LSTMs, combining the input and forget gates into a single update gate and merging the cell state with the hidden state. This makes GRUs computationally more efficient while maintaining similar performance to LSTMs [307].
- **BiRNN:** BiRNN (Bidirectional RNN) combines two RNN layers with opposite information flows, enhancing sensitivity to time sequences. It considers past and future data, making it valuable for dynamic mechanical fault diagnosis, where capturing information across time is vital for identifying fault types and severity. This approach significantly improves fault identification [40].
- **Hybrid Methods:** Moreover, hybrid methods like RNN encoder–decoder (handles variable-length sequences, encoding inputs into embeddings and decoding them to outputs of varying lengths) [37,38,104,308], Convolution RNN (CRNN) models (CNNs extract spatial features and RNNs analyze temporal dependencies) [132,309–311], HMM-RNN [307], and so on are also employed in MFD.

In work on fault diagnosis for reciprocating compressors by [314], researchers optimized LSTM models using Bayesian optimization. The method involved preprocessing vibration signal data from a single sensor to maintain temporal detail while reducing dimensionality. This process, coupled with artificial data augmentation, enhanced the LSTMs' ability to detect 17 distinct fault conditions, outperforming traditional and advanced deep learning techniques. The models' diagnostic accuracy was evaluated using confusion matrices, focusing on metrics like accuracy, precision, recall, and F-score. The best-performing LSTM model achieved a notable 93% average accuracy. Similarly, a deep LSTM (DLSTM) network-based method for predicting the RuL of NASA

[3] AT: Automobile Transmission; B: Bearing; CP: Centrifugal Pump; M: Machinery; PP: Preprocessing; PSO: Particle Swarm Optimization; R: Rotor; UESTC: University of Electronic Science and Technology of China; WTG: Wind Turbine Gearbox.

Table 7
Articles employing RNN in MFD.[4]

AR	M	Dataset	Algorithm(s)	Result	Remarks
[132]	CPF	TEP	CNN (feature learning)-LSTM (time delay capture)	98.88% accuracy	transformed multi-variate time series to 2-D via sliding window
[311]	3-phase IM	Simulation	CRNN	Improved total harmonic distortion	PP: Wavelet transform
[84]	B, G	CWRU, PHM2009	Bi-LSTM for feature denoising and fusion, and a Capsule Network for pattern recognition and fault diagnosis	98.95% diagnostic accuracy with CWRU dataset	PP: normalization and segmentation
[133]	CP	TEP	Bi-GRU for dynamic data processing and cost-sensitive active learning for class imbalance and unlabeled data exploration	96.6% average accuracy in TEP dataset	PP: Batch normalization, L1-norm, and dropout
[312]	B	Proprietary, CWRU	Low-delay lightweight RNN, utilizing LSTM and Just Another NETwork (JANET) cells to reduce RNN"s computational cost	Yielded 1̃00% accuracy with mechanical vibration signals and 9̃5% accuracy using MCSA	PP: signal segmentation
[89]	B (WT)	CWRU, XJTU-SY, NREL	Multi-scale CNN and Bi-LSTM	mean F1 score of 97.12%	feature extraction using multi-scale coarse-grained algorithm, 1-D CNN for learning, BiLSTM for long-term data dependencies, classification with fully connected layer and softmax
[143]	WT-Bl	SCADA	Implemented GSDE model combining CNN with RNN, LSTM, GRU; trained on Chi-square test-based training sets	GSDE model achieved highest overall average rank	Used cost-sensitive learning for DNNs and applied GMDH with cost-sensitive criterion for ensemble predictions
[313]	G (DP)	Proprietary	RNN with simple recurrent units (SRU)	Recognition rate for various faults exceeded 98%	Utilization of SRU and stacked DAE for dimensionality reduction, denoising, and effective handling of complex datasets

turbofan engines, leveraging datasets FD001 and FD003 from the C-MAPSS platform, was developed in the research [39]. The DLSTM model employed multi-sensor data fusion, grid search optimization, and the Adam optimization algorithm, and prevented overfitting with a dropout method. Preprocessing includes noise reduction via exponential smoothing and feature selection based on signal correlation and monotonicity, which led to superior performance indicators such as the lowest Score, R-value, and RUL error range. Again, the study [307] introduces a model for slewing bearing life prediction using an improved GRU network optimized by Moth Flame Optimization, paired with an HMM for early degradation detection. The methodology includes signal preprocessing using Hilbert transform with Robust Local Mean Decomposition and feature extraction in time and frequency domains. This model is verified against standard machine learning methods, demonstrating superior performance in accuracy and robustness for trend prediction and residual life prognosis of excavator slewing bearings. The mean accuracy of 92% was achieved using this model. Furthermore, the study [40] employed a multiple-time window-based CNN bidirectional-LSTM to account for inconsistent lengths of condition monitoring data in the industry and predict the RUL of turbofan engines using the NASA C-MAPSS dataset. The methodology employed data preprocessing that involved feature selection, normalization, and label rectification with a piecewise linear function. Varying time window sizes were used to capture diverse temporal dependencies, enhancing the model's performance over fixed-window methods. A weighted average approach was utilized to aggregate the outcomes from multiple base models, optimizing the ensemble framework's performance.

In the study [315], a DL framework using a BiConvLSTM network, combining the strengths of CNN and LSTM networks was develop for diagnosing faults in planetary gearboxes. Data for the study was collected from a planetary gearbox test rig built by University of New South Wales [316], involving 252 tests across various fault types and conditions. The data, segmented into 2-D matrices, was processed using a bidirectional-convolutional LSTM. This approach outperformed baseline methods like ConvLSTM, CNN-BiLSTM, and others, achieving an overall classification accuracy of 84.72% with 100% accuracy in identifying fault types and locations. Some of the other works employing RNN in MFD are presented in Table 7.

4.2.3. Autoencoder-based FD

Autoencoders are a type of neural network used for unsupervised learning and play a vital role in MFD due to their ability to extract complex features from data and detection of anomalies due to their encoder–decoder structure. AE-based methods are easy to implement and train. They also serve as a nonlinear dimensionality reduction tool by having fewer hidden nodes than input nodes, outperforming kernel PCA.

AEs are crucial for unsupervised learning (USL) in MFD due to their encoder–decoder structure, which is effective in feature extraction and AD. As a nonlinear dimensionality reduction tool, AE outperforms methods like kernel PCA, particularly when configured with fewer hidden than input nodes. They highlight deviations from normal conditions, indicating faults and aiding in the classification of fault types and component diagnostics. This is enhanced by variants such as sparse AE (SpAE), stacked AE (StAE), denoising AE (DAE), variational AE (VAE), and contractive AE (CAE), which handle multimodal and noisy data effectively. Furthermore, incorporating advanced neural modules like CNNs for image data and RNNs for time series allows for dynamic representation learning, making AE a robust choice for ensemble learning to improve the generalization of MFD models [28]. The AE variants employed in MFD are:

- **SpAE:** SpAEs enforce sparsity constraints on the activation of hidden layers, compelling the model to learn a compact and robust representation of the input data. In Sparse Autoencoders, the regularization term is essentially the Kullback–Leibler (KL) divergence, which measures the discrepancy between the distribution of the hidden layer's activations and a predetermined target probability distribution [29]. Some of the studies using SpAE in their work are [90,91,126].

[4] CPF: Chemical Process Faults; DP: Drilling Pump; GMDH: Group Method of Data Handling; GSDE: GMDH-based Selective Deep Ensemble; IM: Induction Motor; WT: Wind Turbine, WT-Bl: Wind Turbine Blade.

- **DAE:** DAEs are designed for robust feature extraction by being trained to reconstruct original, uncorrupted data from inputs deliberately corrupted with artificial noise, thus enhancing the model's resilience to disturbances. Studies [126,318,319] used DAE in their work.
- **VAE:** VAEs are generative models that combine an inference network with a generative network to map input data to a probabilistic latent space. They frame feature extraction as a variational Bayes inference problem, optimizing a likelihood function to effectively learn a probabilistic relationship between observed data and latent representations. This approach imposes a structured probabilistic model on the latent space, where the encoder infers distribution parameters, allowing for robust and meaningful feature generation [28]. [148,186,320–322] are some of the works using VAE in their research.
- **CAE:** CAEs add a penalty term to the classical reconstruction loss function, which helps the model incorporate a regularization term that promotes smooth mapping between input data and latent space. The resulting model exhibits greater robustness to small perturbations in the input data and better generalizes to novel fault patterns. [323–326] are some of the research employing CAE for machinery fault diagnosis using intelligent systems.
- **StAE:** StAEs enhance robust feature extraction by incorporating a regularization term, specifically the squared Frobenius norm of the Jacobian matrix, into the loss function. This term penalizes sensitivity to input variations, thereby encouraging the model to learn smooth mappings from input to latent space and improving its generalization to new patterns, especially in the presence of noise. Before 2018, StAE's general applications were extensively researched [28]. They have been successfully applied to MFD, including, but not limited to [318,320,323].
- **Modified and hybrid versions:** Researchers have proposed various AE modifications and hybrids for improved MFD. Examples include using RNNs or LSTMs with AE to capture temporal dependencies. Convolutional AEs (ConvAE) leverage Convolutional Neural Networks (CNNs) for spatial or temporal data, enhancing performance in fault detection [317,327–329]. Hybrid approaches incorporating clustering or classification algorithms also enhance fault detection and diagnosis systems [330–332]. Some of the works using AE in MFD are presented in Table 8.

4.2.4. Deep belief network-based FD

Deep belief networks, introduced by Hinton et al. in 2006 [333], are a type of DNN constructed hierarchically by stacking multiple layers of restricted Boltzmann machines (RBMs). The first layer of the DBN comprises an RBM that models the input data, while subsequent layers consist of RBMs that capture hidden representations derived from the preceding layers. This architecture facilitates unsupervised pre-training, initializing network weights, and biases using contrastive divergence or persistent contrastive divergence. This pre-training strategy mitigates overfitting and enhances the model's ability to generalize to unfamiliar data. The primary components of RBMs include binary random variables for the visible units and hidden units.

DBNs excel in extracting fault features from specific data representations, and their applications in MFD are growing. The study [95] presented a method for diagnosing bearing faults using DBN. Using 1-D input data from the CWRU bearing dataset, multiple DBNs with different hyperparameters are constructed, which are integrated using an improved ensemble method. The DBNs utilized binary and Gaussian units and were trained using Contrastive Divergence. The method, tested by cross-validation, achieved an accuracy as high as 96.95%, proving its effectiveness in fault diagnosis under challenging conditions. Moreover, in the work [335], an optimized DBN with an improved logistic sigmoidal unit for the fault diagnosis of wind turbine gearboxes is proposed. A dataset of vibration signals from an artificially faulted gearbox is used, and the Morlet wavelet transform, kurtosis index and soft thresholding are used for SP. An improved sigmoidal unit improves convergence speed and classification accuracy: when tested on the MNIST database, for the validation, and gearbox failure data, the

Table 8
Articles employing AE in MFD.[5]

AR	M	Dataset	Algorithm(s)	Result	Remarks
[135]	B,G	THU,UA-FS Gearbox, UO Bearing	Speed normalized AE, normalizing the vibration data to remove the effects of speed vibrations for fault detection	Significant improvement in FD, with highest AUC 0.9704 ±0.0087	PP: low-pass filtering, downsampling, segmentation, and outlier removal
[91]	B, G, TBl	CWRU, MFPT, MaFaulDa, UNSW TBl	Deep sparse AE with Gray Wolf Optimization	100% accuracy in other datasets, 95% on MaFaulDa	Rprop for training
[126]	B, G	PU, Qianpeng Company	Denoising integrated sparse AE: feature enhancement and denoising based on fault sensitivity, data decoupling, and adaptive loss function	Over 99% test accuracy	PP: Ensemble empirical mode Decomposition; Emphasis on dealing with weak signals and background noise
[317]	B	Proprietary	Bayesian optimization and channel-fusion-based ConvAE	more than 99% of accuracy, precision, recall and f1 score	PP: Vibration signals segmented using a sliding window technique
[147]	B	XJTU-SY	Deep spatiotemporal fusion AE network, incorporating multiscale convolution, convolutional LSTM, and attention mechanism; Integration of multimode samples for unsupervised health indicator construction, introducing a quadratic function-based shape constraint	Average comprehensive score of 0.7327 on bearing dataset	PP: Wavelet threshold denoising and Sparrow Search Algorithm-Variational Mode decomposition; Input Method: Coupling of vibration signal, vibration frequency domain signal, Intrinsic Mode Function, and its frequency domain signal into a 2-D matrix, forming multimode-coupled samples
[148]	B	XJTU-SY, Proprietary	Deep order-wavelet convolutional VAE, integrating improved energy-order analysis and Wavelet Kernel Convolutional Block; Employed frequency-weighted energy operator, anti-symmetric real Laplace wavelet, and multi-objective gray wolf optimizer	Average identification accuracy 99%	Improved energy-order analysis for transforming time-domain vibration signals into angle-domain signals

[e] Rprop: Resilient Backpropagation; TBl: Turbine Blade.

Table 9
Articles employing DBN in MFD.[6]

AR	M	Dataset	Algorithm(s)	Result	Remarks
[94]	B	CWRU	Self-adaptive DBN, optimized using the Salp Swarm Algorithm	overall classification accuracy of 94.4%	Extract the time domain, frequency domain, and time-frequency domain features
[127]	B	PU	Mixed pooling DBN	98.84% accuracy	Morlet wavelets for generating time-frequency images
[92]	B	CWRU, Proprietary	Bi-directional DBN with forward training for feature learning and reverse generation for sample synthesis; Quantum Genetic Algorithm for parameter optimization	Average accuracy percentages: 96.57%, 94.58%, 93.74% for different imbalance ratios	PP: Normalization and truncation of data; noise time-shift layer added to reduce sample similarity; DataAug: Reverse generation part of Bi-DBN synthesizes supplementary samples to address imbalanced datasets
[134]	CP	TEP	DBN with extended RBMs, integrating features from raw data for layer-wise feature extraction and fault classification in chemical processes	Best average accuracy of 94.31%	DataAug: Dynamic data augmentation to capture temporal correlations
[93]	B	CWRU, QPZZ	CWT for transforming raw data to RGB images and employing Gaussian Convolutional DBN for fault classification	Average accuracy of 99.57%	PP: Scalogram construction from vibration signals, RGB image transformation, and whitening processing of images
[334]	G	Simulation	DBN trained with labeled gear fault signals, Sparrow Search Algorithm for optimizing DBN parameters	Average detection accuracy of 96.18%	PP: 13 time-domain features are extracted from each sample group of raw data
[97]	B	CWRU, Hydraulic Pump	Hilbert envelope spectrum and DBN	Classification accuracy of 99.55%	PP: Resampling of vibration signals, application of anti-aliasing filter
[144]	WTB	SCADA dataset	DBN with back-propagation and layer-wise training, Exponentially weighted moving average for monitoring prediction errors; binary vectors for fault classification	AUC of 0.88 obtained with LS-SVM classifier	Parameters reduced from 33 to 12 using domain expertise; wrapper with genetic search, boosting-tree algorithm, and relief algorithms were used for final selection in generator bearing temperature prediction

method outperforms conventional units, achieving high accuracy up to 96.32% and fast convergence. Furthermore, Yan et al. in [96] proposed the multiscale cascaded DBN for fault detection in rotating machinery using raw vibration signals from the CWRU bearing dataset and North China Electric Power University gear vibration dataset. Employing an improved coarse-grained multiscale process and a three-layer DBN for feature extraction, the proposed method used data segmentation and Fourier spectrum calculations as preprocessing techniques, obtaining over 99% classification accuracy.

Another study by Tao et al. [336] focused on bearing fault diagnosis using a DBN architecture, consisting of multiple layers of RBM and backpropagation neural network (BPNN), and multisensor information fusion. The method involved acquiring multiple vibration signals from various faulty bearings, extracting time-domain characteristics from these signals, and then inputting the data into the DBN to generate a classifier for fault diagnosis. Experiments were carried out on the QPZ-II dataset, and the comparisons were made with SVM, KNN, and BPNN methods, demonstrating that the DBN-based method achieves higher identification accuracy of 97.5% for training samples and 95.5% for testing samples. Moreover, the study [337] focused on using three different architectures: deep Boltzmann machine, DBM and StAE. The method involved the pre-processing of vibration signals through four schemes covering the time, frequency, and time-frequency domains. The dataset includes seven failure patterns of rolling bearings collected from rotating machinery systems. The performance of the DNN model is evaluated in terms of accuracy, which is more than 99% in the best setting.

Apart from the classification, DBN is used in predicting RUL of rotating machinery. The study [106] employed DBN, combining with local linear embedding (LLE), and difussion process for predicting RUL of bearings of FEMTO dataset. DBN was used for feature extraction and LLE for health index (HI) determination. HI evolved based on difussion process and a probability density function of the predicted RUL was derived in terms of the first hitting time. Moreover, a study [338] proposed an unsupervised learning-based fault diagnosis model for rotating machinery, integrating the SpAE, DBN, and binary processor. The method used SpAE to encode signals in the frequency domain, which were then processed by a binary processor and fed into a DBN for fault diagnosis. The process did not require labeled training data making it entirely unsupervised. The dataset used was the CWRU bearing dataset and gear pitting dataset, from which time domain signals were transformed into normalized frequency domain signals using FFT. During preprocessing, a binary processor converted the SpAE outputs to binary data, thereby improving the RBM's efficiency within the DBN. Some of the other works employing DBN in MFD are summarised in Table 9.

4.2.5. Generative Neural Network-based FD

A Generative Adversarial Network (GAN), introduced by [339] in 2014, is an outstanding unsupervised generative algorithm that learns to create realistic data from a random distribution. GAN, which is acknowledged as one of the 'Top Ten Global Breakthrough Technologies' [26], consists of a generator (G) that produces synthetic samples from random noise and a discriminator (D) that distinguishes real instances from synthetic ones. Both the generator and discriminator are designed as deep neural networks. The generator maps a latent variable to the data space using its parameters. The discriminator estimates the probability that a sample is genuine or fake using its parameters. This setup forms a two-player minimax game.

GANs have been used in many research fields, including natural language processing, computer vision, and so on. They were initially used in fault diagnosis as for data augmentation, a strategy to generate additional instances with the same data distribution to solve the problem

[6] AUC: Area Under the Curve; DataAug: Data Augmentation; LS-SVM: Least Square Support Vector Machine; WTB: Wind Turbine Bearing.

of small sample size. Then, GANs were adapted to adversarial cross-domain fault diagnosis, known as adversarial domain adaptation, which differs from data augmentation as it uses adversarial training with both target and source domain data to extract domain-invariant features. Moreover, these algorithms have also been applied to semi-supervised learning and anomaly detection [121,340]. Various improvisations and improvements have yielded different variants of GANs, which have been developed and practiced to solve the existing limitations, like model collapse and imbalance training. The variants can be categorized into two categories according to the improvements made, and are called structure-focused improvements and loss-focused improvements [26].

I. **Structure-focused Improvements:** GANs, based on the improvements in their structure are further divided into three categories: convolution-based, condition-based, and semi-supervised GANs. To address the original GAN's feature extraction and training inefficiencies, the deep convolutional GAN (DCGAN) was developed [184]. It employs convolutional and deconvolutional layers in the discriminator and generator, improving stability and utilizing weight sharing and local connections for enhanced performance. On the other hand, condition-based GANs, including the conditional GAN (CGAN) [181], InfoGAN [341], and Auxiliary Classifier GAN (ACGAN) [342], tackle standard GANs' mode collapse issue. CGAN uses class information for guided generation, InfoGAN employs a latent code and an extra classifier to enhance input–output correlation, and ACGAN integrates an auxiliary network for classifying faults and distinguishing real from synthetic data. Furthermore, the Semi-Supervised GAN (SSGAN) efficiently uses unlabeled data in low-labeled-data scenarios. It features a softmax classifier in the discriminator for distinguishing real from synthetic inputs and classifying real samples, enabling semi-supervised learning (USL) [341].
II. **Loss-focused Improvements:** Loss-focused GANs were developed to stabilize training and address issues like unstable gradients and mode collapse in standard GANs. Wasserstein GAN (WGAN) [343] uses Wasserstein distance for a more stable divergence measurement between real and fake samples, though limited by weight clipping. The WGAN with Gradient Penalty [185] further refines this by adding a gradient penalty to address these limitations. The Least Squares GAN (LSGAN) [344], proposed by Mao, employs least squares loss functions to stabilize training by penalizing samples based on their proximity to the decision boundary. Additionally, the Energy-Based GAN (EBGAN) [345] and Boundary Equilibrium GAN (BEGAN) [346] innovate with autoencoder and encoder–decoder structures in their discriminators, respectively, the latter featuring an equilibrium enforcing algorithm. These variants collectively enhance the standard GAN framework by focusing on loss function and training strategy modifications to resolve specific challenges.

The application of GANs, also tabulated in Table 10, in MFD can be classified as:

- **Data Augmentation and Balancing:** Recent advances in GAN-based data augmentation and class balance help to solve the problem of small samples and data imbalance in fault diagnosis. For the data augmentation, the standard procedure is to collect various fault state data, train the GAN with real instances, and train a classifier by combining the generated synthetic and real data. GANs are mainly used for mechanical fault diagnosis, especially with vibration signals collected by sensors, and from limited data, fake samples are generated. These methods can be classified into 1-D time domain, 1-D frequency domain, 2-D image signals, and 1-D feature sets for generating synthetic data [26,347].
- **Anomaly Detection:** The AD in GAN-based MFD is becoming increasingly important, especially when only normal operating data are available. The method relies on learning with normal samples to draw boundaries that distinguish between normal and faulty states. Modern approaches use adversarial learning to improve this process and move away from manually generated features. The core idea is to use GANs to generate synthetic samples and use the reconstruction loss between these samples and the original normal samples to detect anomalies. This approach has gained acceptance not only in FD but also in other areas requiring efficient and reliable AD [130,340].
- **Semi-supervised Adversarial Learning:** In situations where labeled data is scarce, semi-supervised learning is employed to utilize unlabeled data for model training. This approach combines unlabeled data with adversarial strategies to enhance training [65,121].
- **Adversarial Training for Transfer Learning:** Adversarial training is applied to transfer learning, called adversarial domain adaptation (ADA), which uses source domain data to enrich limited target domain data. ADA models fall into two types: adversarial discriminative models, which create domain-invariant features for fault classification, and adversarial generative models, which facilitate domain adaptation by learning data distributions or transforming data between domains [67,68].

Table 10
Application of GANs in MFD.

GAN as:	Article references
Data Augmentation and Balancing	[56,59–63,181–185,187,348–357]
Anomaly Detection	[64,109,130,155,358–362]
Semi-supervised adversarial learning	[65,66,110,121,156]
Adversarial Training for Transfer Learning	[67–69,157]

4.3. Reinforcement learning-based FD

Reinforcement learning, a sub-field of ML, is a computational technique that focuses on training agents to make decisions in an environment by interacting with it and learning from the feedback received as rewards or penalties. RL has its roots in two primary research domains: the first being optimal control through the utilization of value functions and dynamic programming, and the second drawing inspiration from animal psychology, particularly the concept of trial-and-error search [363]. RL distinguishes itself from supervised learning (SL) in the context of connectionism [364]. In RL, the feedback signal received from the environment assesses the action's effectiveness, rather than instructing the system on how to produce the correct action [149]. The primary goal of an RL agent is to learn an optimal policy that maps states to actions, maximizing the expected cumulative reward over time. This learning process can be achieved through various algorithms, such as Q-learning, Deep Q-Networks (DQN), Proximal Policy Optimization (PPO), Actor-Critic methods, and so on [365].

In recent years, there has been a growing interest in applying RL to machinery fault detection, diagnosis, classification, and RUL prediction. RL has found applications in diverse areas, including fault diagnosis in transmission lines [366], optimizing operations in the smart grid [367], fault detection in Hydraulic press [368], industrial process control [363]. Moreover, the use of RL can be observed in both offline and online scenarios. In the offline context, researchers have explored how RL can be leveraged for manufacturing industries, as evidenced by studies [369,370]. Offline RL often involves training a model on historical data to make predictions and decisions [371]. On the other hand, RL is also applicable in online scenarios, where it operates in real-time to identify faults and take corrective actions promptly. The utilization of RL in online settings is a dynamic and evolving area. On the other hand, RL is also applicable in online scenarios, where it operates in real-time to identify faults and take corrective actions promptly. The utilization of RL in online settings is a dynamic and evolving area. Studies [117,372] provide insights into this aspect. The choice between offline and online RL in machinery fault detection depends on factors such as the nature of the machinery, the availability of real-time data, and the desired response time. Both approaches offer valuable contributions to the field and are continually evolving to enhance industrial processes, improve equipment reliability, reduce downtime, and ultimately advance the overall efficiency and safety of machinery operations.

The utilization of reinforcement learning in machinery fault detection, diagnosis, and classification involves training RL agents to make decisions based on the observed data, such as vibration signals, acoustic emissions, or thermal images. The RL agent can learn to identify different fault conditions by interacting with the data and receiving feedback in the form of rewards or penalties. In most of the MFD studies, fault diagnosis is approached as a classification task resembling a guessing game. The researchers created a simulation environment resembling a 'fault diagnosis game'. It presents questions with fault samples and labels to the agent, which must diagnose them. In a K-class fault diagnosis scenario, the guessing action space ranges from 0 to K-1, with 0 denoting normal and k representing the Kth fault type. The agent receives rewards for accurate guesses and penalties for incorrect ones. Over multiple rounds of this guessing game, the agent aims to learn an optimal policy for fault identification using sensor data from monitoring equipment. A similar approach is also seen in [70], in which the researchers employed the RL method for intelligent fault diagnosis for rotating machinery. The agent, constructed using a stacked autoencoder, learns fault diagnosis through a DQN. The method combines RL and DL, enabling end-to-end fault diagnosis for machinery, in which Memory replay and rewards help the agent learn fault mappings from raw vibration signals with minimal external guidance.

Other than the classification task, the researchers proposed an automatic neural architecture search (NAS) approach using reinforcement learning in [74]. This study employed a special RNN unit called *Nascell* within a generator network and a controller layer composed of two stacked Nascell units. The controller outputs parameters for constructing CNN architectures, including convolutional kernel sizes, kernel numbers, and pooling sizes for each layer. The process involves iteratively generating network configurations, training child models, and maximizing model accuracy as the reward through policy gradient updates. This iterative approach continues until the search converges to an optimal architecture, occasionally injecting randomness to prevent getting stuck in local optima. In short, the controller acts as an agent, shaping the CNN architecture, and reinforcement learning is used to navigate the architectural search space toward higher-performing models. Moreover, In a study [159], multi-label transfer reinforcement learning (ML-TRL), for compound fault diagnosis in bearings, is employed. It combines deep RL (DRL) and TL to enhance fault feature extraction and improve accuracy in recognizing compound faults. ML-TRL outperforms traditional methods, and its use of transfer learning involves pre-training convolutional layers, reducing the complexity of DRL training.

Similarly, [72] employed an automated approach for designing fault diagnosis models using a combination of RL and NAS. Notable optimizations include a Greedy Strategy to prevent local optima, ER to smooth the learning process, and Weight Sharing to reduce computational demands. Moreover, the researchers employed a DQN-based RL framework to optimize a bandpass filter's upper and lower cutoff frequencies for fault diagnosis in rotating machinery signals in [149]. The bandpass filter acts as an agent, and its position, defined by these frequencies, represents its state. The agent interacts with the signal environment to maximize a reward signal based on how effectively it highlights fault characteristic frequencies. The DQN algorithm iteratively refines the frequency band to reveal the optimal range for fault identification. Experimental results on gear and bearing fault signals demonstrate that the proposed deep reinforcement learning-based method outperforms other traditional techniques, such as fast Kurtogram and GiniIndexgram, in identifying fault-related frequencies.

Study [117] focused on developing an online domain adaptation learning method for fault diagnosis in machinery, which is achieved through the use of a Capsule Network (Cap-net) as an agent for autonomously extracting fault features from online data. A feature dictionary based on Coarse-grained Similarity (CS) is designed to label the online data, and a reward mechanism based on CS is used to evaluate the coarse-grained labels. The method consists of initializing the Cap-net, updating it with online data responses and rewards, and fine-tuning it with historical data. The target and evaluation networks are updated iteratively, and a self-pruning mechanism is employed to optimize the online feature dictionary. Some of the other works employing RL in MFD are mentioned in Table 11.

4.4. Transfer learning-based FD

Transfer learning, a promising paradigm in ML, addresses the limitations of DL in practical applications, particularly when dealing with limited data and distribution discrepancies between training and testing datasets [21]. DL models typically require large amounts of labeled data to perform well, which is often impractical in industrial settings where collecting such data, especially faulty data, is time-consuming and challenging. TL mitigates this issue by transferring knowledge from related but different domains to improve model performance in new tasks, thus addressing data scarcity and reducing computational demands [373]. By utlizing knowledge from related domains, TL not only enhances learning with limited data but also ensures robust generalization and effective feature extraction from complex datasets. This makes TL a valuable tool in applications where traditional ML methods may struggle. In MFD, TL is widely applied, and for this review, we have categorized its use into two categories:

4.4.1. Direct transfer

The direct transfer approach involves using pre-trained models developed on one dataset (source domain) and applying them to another dataset (target domain) with minimal or no additional training. This method is particularly useful when data from the target domain is scarce, as it allows the model to use patterns and features learned from similar but different conditions. The success of this approach depends on the similarity between the source and target domains; more similar domains lead to better performance. Although fine-tuning may be needed if the domains differ significantly, the direct transfer approach can effectively address the issue of labeled data scarcity. For instance, a model pre-trained on a large, labeled dataset like the CWRU bearing dataset can be applied to a target domain with sparse data or different operating conditions. The model's performance is evaluated using metrics like accuracy and precision, and fine-tuning is conducted if necessary to ensure it adapts well to the new domain. Studies [46–48,114,374–376] are some of the works employing pre-trained models for MFD.

4.4.2. Domain adaptation

Recently, domain adaptation methods have been utilized widely in MFD, particularly when dealing with scenarios when there is a discrepancy between the source and target domains, Unlike direct transfer methods, which may suffer from significant performance degradation due to domain shift, DA methods are designed to mitigate these issues by focusing on reducing the distribution differences between domains. Domain adaptation relaxes the conventional assumption that training and testing data must be independent and identically distributed *(iid)*. By utilizing invariant features and essential structures across different but related domains, DA methods effectively address challenges such as domain shift, small sample sizes in the target domain, and unbalanced datasets [21,373]. In this approach, a previously trained model is fitted to a new but related domain to minimize discrepancies between domains. The use of DA methods in deep TL can be subdivided

Table 11
Articles employing RL in MFD.[7]

AR	M	Dataset	Algorithm(s)	Result	Remarks
[71]	B, G	CWRU, UoC	FD as a guessing game; Classification using 1-D CNN-RL and GRU-RL	CNN-RL: CWRU=99.95%, UoC=99.61%; GRU-RL: CWRU=99.98%, UoC=99.95%	Actions are chosen via ϵ-greedy method; ER updates Policy-Net
[131]	G	SEU	FD as a guessing game; DQN agent uses two 2-D CNNs with the same structure: Eval-Net (Updating) current Q value, Target-Net (Calculating fixed target Q value)	Overall accuracy > 99.5%	raw signals transformed into 2-D TF maps (agent's observation) using synchro-extracting transform; ϵ-greedy exploration
[41]	B, G	CWRU, THU	FD as a guessing game; DDQN for solving class imbalance problem	Increment in class imbalance decreased the F1 scores	Higher rewards for diagnosing minority classes
[72]	B, G, BS	CWRU, PHM-2009, IMS, HOUDE, BS	Automatic NAS using RL; Creates CNN models for fault diagnosis tasks; Greedy Strategy to avoid local optima	100% accuracy in test data	ER for smoother learning; Weight Sharing to reduce computational demands
[149]	B, G	MFS-SQ	DQN-based RL optimizes bandpass filter cutoff frequencies for FD; Bandpass filter acts as an agent, with frequencies defining its state.	Outperforms fast kurtogram, GiniIndexgram, and Smoothnessgram methods	DQN iteratively refines the frequency band for optimal fault identification
[117]	B	PU, MFPT	Utilized a Capsule Network for online domain adaptation; A reward mechanism based on CS evaluates labels; Self-pruning mechanism optimizes the online feature dictionary	Accuracy for different testing conditions are MFPT: Average accuracy 92% and PU: ≥ 96.25%	T-SNE for feature visualization in 2-D space;target and evaluation networks are updated iteratively
[115]	B	PU, JNU	Contrastive augmented DRL for fault classification in class imbalanced data; Two-stage training: Pretraining (CL to distinguish fault classes), fine-tuning (dynamic reward function for improved recognition of minority class)	6%–7% accuracy improvement than other models	Adaptive rewards improved stability & performance; Contrastive loss minimized overlap

into three categories: the discrepancy-based method, the adversarial-based method, and the reconstruction-based method, which are briefly described below.

1. **[A.] Discrepancy-based Methods:** The discrepancy-based DA methods aim to minimize the differences between the source and target domains by reducing the discrepancy in the feature layer within a neural network. This method primarily focuses on aligning the feature distributions of both domains, enabling the model to generalize effectively to the target domain. In discrepancy-based DA, the core idea is to measure and reduce the distance between the source and target domains by applying statistical approaches to the feature layers of the model [373,377]. Statistic transformation, structure optimization, and geometric transformation are key strategies for minimizing discrepancies between source and target domains in domain adaptation. By adjusting the statistical properties of feature distributions, such as mean and variance, techniques like Maximum Mean Discrepancy (MMD) and Correlation Alignment (CORAL) reduce domain differences. Structure optimization involves modifying the model's architecture or adding layers to enhance domain alignment. Additionally, geometric transformation aligns the feature space's geometric properties, ensuring consistency between the source and target domains, further reducing discrepancies. Studies [49,50,55,57] implemented discrepancy-based DA approach in their work.
2. **[B.] Adversarial-based Methods:** Adversarial-based DA methods employ a domain discriminator to promote domain confusion and learn invariant features between source and target domains. Inspired by a two-player game as in GAN, this approach consists of a generator and a discriminator competing against each other, where a generator tries to create data that closely resembles the target domain, while the discriminator attempts to distinguish between the source and generated data. The adversarial training process involves the generator trying to confuse the discriminator, leading to better domain alignment. These methods are divided into Generative Adversarial DA (GADA), a method with an additional generator, and non-generative adversarial DA (non-GADA), a method without an additional generator. GADA utilizes a generator to produce data samples that resemble the target domain, aiding in domain alignment through transformations like those seen in Coupled GANs and adversarial discriminative domain adaptation (ADDA). Non-GADA, however, focuses on domain alignment without generating new samples, using the generator as a feature extractor to maximize domain confusion. Techniques like Wasserstein distance-based optimization, curriculum learning, and domain-symmetric networks are employed to reduce domain shift. While adversarial-based DA methods offer advantages like high sample diversity and theoretical proximity to target data, they face challenges with training stability due to difficulties in achieving Nash equilibrium [24,373]. Studies [51,52,157,378–381] employed adversarial-based DA method in their work.
3. **[C.] Reconstruction-based Methods:** Reconstruction-based DA method focuses on reducing domain differences by reconstructing the data from both source and target domains, ensuring intra-domain distinctions while capturing inter-domain commonalities. These methods typically use a coding-decoding framework, where data is encoded into feature representations and then decoded back to the original input, creating a shared domain space for alignment. By sharing an encoder, the model learns domain-invariant features while preserving domain-

[7] BS:Ball Screw; G:Gears; LBD:Locomotive Bearing Dataset; TF:time-frequency.

specific characteristics. Techniques like self-encoders and KL divergence enhance domain alignment, as seen in domain separation networks, which separate domain-invariant and domain-specific features to prevent negative transfer. Hybrid models combining AEs with GANs further improve performance by incorporating tasks like cyclic consistency loss, which is particularly useful in unsupervised scenarios. Despite the advantages, challenges remain in balancing the authenticity and diversity of generated samples and managing the complexity of parameter updates, making reconstruction-based DA a powerful yet intricate approach to domain alignment [21,373]. Studies [53,125, 152,382] applied reconstruction-based techniques in their work.

4.4.3. Multi-source domain adaptation

Multi-source domain adaptation (MDA) uses data from multiple source domains with different distributions to improve model performance on a target domain. Unlike traditional single-domain adaptation, MDA addresses the challenge of conflicting information by aligning features across multiple domains, enhancing the model's generalization. Key approaches include residual adaptive modules for compressing and sharing parameters, latent domain discovery for finding commonalities, and domain weighting to select the most relevant source data. Adversarial training with GANs further strengthens MDA by minimizing distribution differences between domains. Despite its effectiveness, MDA still faces challenges in automatically selecting and aligning the most relevant data, making it a focal point for ongoing research in fault diagnosis [2,373].

4.4.4. Partial transfer learning

Partial transfer learning is a specialized approach where only a subset of features or tasks from the source domain is relevant to the target domain [21]. In the context of MFD, this involves selectively transferring knowledge related to specific faults or operational conditions, rather than applying the entire model from the source domain. This method is particularly useful when the target domain has a narrower scope or when the source domain contains redundant or irrelevant information. By focusing only on the most relevant data, partial transfer learning minimizes the risk of negative transfer — where irrelevant knowledge could degrade model performance — and ensures that the transferred information enhances the target model's accuracy and effectiveness. However, identifying the relevant subset requires careful analysis and domain expertise, making it a challenging but crucial aspect of successful fault diagnosis [382].

With the concepts of TL outlined, these approaches have been increasingly applied to address practical challenges in MFD. Guo et al. [99] addressed the challenge of limited labeled data and differing training/testing distributions by employing a deep convolutional TL network. This network comprises two modules: condition recognition and domain adaptation. The condition recognition module uses a 1-D CNN to automatically extract features from raw vibration signals and classify machine health conditions. The domain adaptation module further enhances the 1-D CNN by incorporating a domain classifier and a distribution discrepancy metric, enabling the model to learn domain-invariant features. This setup maximizes domain classification errors while minimizing the probability distribution distance between source and target domains. The approach was validated through six transfer fault diagnosis experiments on three different bearing datasets, showing significant improvements in fault diagnosis accuracy compared to traditional methods. Similarly, the study by Wei et al. [136] presented a method for machinery fault diagnosis of THU planetary gearbox dataset and UA-FS gearbox dataset. The methodology involved using the raw vibration signals and weighted domain adaptation network to address the challenge of data distribution shifts due to changing working conditions by assigning weights to conditions based on their similarity target. The study demonstrates improved diagnosis accuracy, utilizing classification accuracy and MMD as metrics, emphasizing the importance of domain adaptation in variable working conditions. Another study [80], proposed the wavelet packet transform (WPT)-based deep feature transfer learning method for bearing fault diagnosis under different working conditions, in which the combination of WPT (for time–frequency feature map construction), a deep ResNet (for features extraction), and a multi-kernel MMD (for evaluation of distribution differences in features across domains) is used. The study carried out on the CWRU bearing dataset, MFS-RDS (Rotor Dynamics Simulator) dataset provided average of 08.59% accuracy result in the CWRU dataset, whereas 97.14% accuracy in the MFS-RDS dataset. Talking about multi-source domain adaptation, the study [68] by Rezaeianjouybari et al. highlighted the limitations of previous models that rely on a single source domain and ignore variations in working conditions, and proposed feature-level and task-specific distribution alignment multi-source domain adaptation (FTD-MSDA) model to address the challenge of domain shift in intelligent fault diagnosis systems. Experimented on the CWRU and PU bearing dataset, the FTD-MSDA model framework aligned domains at both the feature and task levels, using Sliced Wasserstein discrepancy for shaping task-specific decision boundaries and successfully transferred knowledge from multiple labeled source domains to a single unlabeled target domain. Moreover, a study by Qian et al. [383], introduced the TL method called adaptive intermediate class-wise distribution alignment, targetting the fault diagnosis in the wind turbine planetary gearboxes, addressing the challenges like slow convergence and loss oscillation in DA. The proposed model, employing both domain adaptation and generalization, utilized an adaptive intermediate distribution mechanism alongside an AdaSoftmax loss, which dynamically aligns the source and target domain distributions without requiring additional losses for distribution distance or correlation regularization. Furthermore, a study by [384] developed the variance discrepancy representation method to enhance the DA in FD of rotating machinery. The experimental results across three bearing datasets demonstrated that the proposed method improved the representation of distribution discrepancies by focusing on variance rather than mean, providing a more accurate reflection of differences between source and target domains. Some of the other works employing transfer learning are presented in Table 12.

4.5. Other methods

4.5.1. MFD as anomaly detection

In the field of MFD, where the lack of labeled fault data is a challenge, un/semi-supervised AD techniques play an important role in identifying faulty patterns and potential anomalies that cannot be captured by traditional supervised methods. These data-driven approaches focus on detecting outliers or anomalies (data points that deviate significantly from the majority of normal data instances) [390]. Various classical unsupervised methods are used to identify anomalies in datasets. These include Z-Score, which measures how many standard deviations a data point is away from the mean; Interquartile Range, which identifies outliers by considering the range between the first and third quartiles of the data; Isolation Forest (iF), a tree-based algorithm that isolates anomalies by constructing isolation trees and measuring the number of splits required to isolate a data point; Local Outlier Factor (LOF), which measures the local density deviation of a data point with respect to its neighbors; and One-Class SVM, which learns a boundary around normal data instances and classifies anything outside this boundary as an anomaly, and so on. Apart from these shallow learning algorithms, DL-based approaches, such as AE, VAE, and GAN, have also been used for AD in different ways, such as feature extractors, representation learners from normal data, and end-to-end anomaly score learners. More details can be found in [390].

A study by [391] introduced the full graph dynamic AE (FGDAE)-based AD method, designed to operate effectively under complex and varied conditions. The FGDAE model integrated a fully connected graph

Table 12
Articles employing TL in MFD.[8]

AR	M	Dataset	Algorithm(s)	Result	Remarks
[90]	B	CWRU, PU	Transferable decoupling multi-scale AE: Multi-scale residual network with transposed convolution for distribution alignment, feature extraction, classification, and domain adaptation	Achieved upto 100% in domain transfer tasks in certain sub-datasets	PP: Data cleaning and segmentation
[385]	B	QPZZ-II	Joint distribution adaptive DBN optimized with a Sparrow Search Algorithm variant	Average accuracy 79.21%, peak accuracy 90.71% in specific conditions	PP: Time-domain data converted to frequency-domain via FFT
[386]	G	WT-DDS	Deep TL integrating cohesion-based sensitive feature selection, a three-layer SpAE for feature extraction, MMD for aligning data distributions across different datasets, and SoftMax as classifier	Highest classification accuracy of 99.17%	Signals from four sensors averaged over three experiments
[387]	B, G	CWRU, Gear Fault Dataset	Integration of unsupervised geodesic flow kernel (GFK)-based DA with Z-normalization and SoftMax regression	Average detection of 99.9%	PP: STFT to raw vibration data, converting it to a power spectral density matrix in decibel form, and then flattening it into a 1-D vector for GFK algorithm compatibility
[98]	B	CWRU, Self-built	Deep imbalanced domain adaptation, integrating empirical risk minimization, cost-sensitive re-weighting (class-balanced loss), categorical alignment (local MMD), and margin loss regularization (label-distribution-aware margin loss)	Average accuracy of 96.55%	Multi-channel vibration signal as input to 1-D ResNet, used as backbone
[388]	B, WTG	CWRU, WF Dataset	Instance-based deep boosted TL, utilizing algorithms like AdaBoost and TrAdaBoost for weight adjustments to minimize negative transfer effects	9.79% improvement from base case	PP: Noise removal; DataAug:Sliding window technique
[389]	B	CWRU, self-built	Self-attention lightweight CNN and TL	Significant reduction in model parameter	PP: CWT

to capture global structural relationships between sensor channels, a graph adaptive AE that aggregates multi-perspective features and adapts to changes in operating conditions, and a dynamic weight optimization strategy to handle training with unbalanced multi-condition data. Similarly, a study [130] proposed fault-attention generative probabilistic adversarial AE (FGPAA) method for AD of three machinery fault datasets by focusing only on healthy classes. Utilizing dual adversarial AEs, the FGPAA method employed a fault-attention probability distribution to evaluate the health state of machinery effectively, allowing for dynamic adjustment to signal noise and anomaly detection in real time. Moreover, a work by [392] proposed a real-time AD method for the fault diagnosis of marine machinery. The authors develop a framework named RADIS, based on LSTM-based VAE combined with multi-level Otsu's image thresholding technique. Furthermore, a GAN-based AD method was proposed in [358], where an encoder–decoder-encoder architecture was employed in the generator to train the normal samples exclusively. The latent and apparent losses to compute anomaly scores.

4.5.2. Use of transformers

The Transformer, initially proposed by Vaswani et al. [393], revolutionized natural language processing and has been successfully applied in other areas like computer vision. These models, characterized by layers of Transformer blocks featuring multi-head self-attention and batch normalization, efficiently handle tasks without complex operations like CNNs or RNNs, often outperforming them. Recently, the use of architectures has been seen in MFD as well. A study [79] introduced a time-series transformer (TST) model for direct 1-D raw vibration data processing, without any signal preprocessing. TST leverages multi-head self-attention and transformer blocks for feature extraction from bearings and gear faults. Evaluation on the CWRU, XJTU, and UoC datasets shows impressive accuracy, e.g., 98.63% for CWRU and 99.78% for XJTU and 99.51% for UoC. TST's feature vectors exhibit superior intra-class compactness and inter-class separability, emphasized by t-SNE visualization. Similarly, a window-based multi-head self-attention model is proposed in [78], employing three datasets: CWRU, UoC, and Shandong University (SDU) dataset. Data preprocessing includes 1024 sample lengths and dataset partitioning. The model combines self-attention and CNN with a 1-D window-based multi-head self-attention for local feature learning. Results demonstrate the model's superior classification performance, achieving nearly 99.99% accuracy in noise-free conditions and maintaining robustness with added noise (SNRs from −6 dB to 6 dB). Moreover, an article by Wu et al. introduced a transformer-based classifier for machinery fault classification [394]. Employing CWT for the generation of time-frequency spectrogram images from raw data as an input, the authors employed a Mahalanobis distance-based technique to identify previously unseen faults.

4.5.3. Physics informed neural networks (PINN)

PINNs are an innovative approach in ML, which incorporates physical laws into the structure of neural networks. This integration enhances predictive accuracy and interpretability. In the domain of MFD, PINNs offer a noteworthy advancement by enabling precise identification and analysis of faults in mechanical systems. A study [395] employed PINN, applying physics-informed loss functions, to enhance the model's interpretability for fault severity identification of axial piston pumps. The research utilized a high sampling rate to collect data on axial piston pumps, identifying piston wear in four severity levels using a low-pass filter to isolate relevant frequencies. The proposed PINN model accurately identified wear states by estimating gap clearances linked to the pump's health indicators. Similarly, researchers studied early fatigue in the main bearings of wind turbines and aimed to predict the RUL of these bearings using PINN method in [396]. This model assessed bearing fatigue and grease degradation using data such as wind speed, bearing temperature, and grease analysis from a 1.5

[8] WF: Wind farm; WT-DDS: Wind Turbine Drivetrain Diagnostics Simulator; WTG: Wind turbine gearbox

MW wind turbine. Moreover, Ni et al. [397] implemented a physics-informed residual network for fault diagnosis using a rolling element bearing dataset from a drivetrain diagnostics simulator. The network featured a modal-property-dominant layer with cepstrum exponential filtering to emphasize system properties, a domain-conversion layer using Computed Order Tracking to address speed variations, and a parallel bi-channel architecture for extracting complex fault features, resulting in improved diagnostic accuracy under varying conditions.

4.5.4. Federated Learning

Federated Learning introduces a decentralized model training method, especially useful in edge learning environments where data is distributed across multiple devices. It trains a shared model via iterative updates using local data, avoiding the centralization of sensitive information. This technique facilitates collaborative training while safeguarding data privacy and minimizing communication overhead. FL has been widely adopted for edge ML models and successfully applied in areas like cyber attack detection, spam detection, smart cities, autonomous vehicles, and recently machinery fault diagnosis, primarily aiming to maintain privacy [398]. In a study by [399], the researchers implemented an FL-based method for detecting mixed faults, particularly focusing on the rotor and bearing. The input method is 1-D time-series data collected from accelerometers, with a dataset comprising 92,160 points across 48 mixed fault classes. The methodology includes an FL approach utilizing a duplet classifier, where data is distributed among 30 clients in three partition schemes: balanced *iid*, balanced non-*iid*, and unbalanced non-*iid*. Preprocessing involves data shuffling and partitioning. The FL model, trained with CNNs, achieved over 90% classification accuracy for mixed faults. Another study by [400] focuses on improving fault diagnosis in rolling bearings using federated learning. Using data from the CWRU benchmark dataset and a project dataset from Nio Inc., the study aims to address client inhomogeneity issues (related to sample size, quality, and fault type) in FL. The proposed methodology is a multi-scale layer-by-layer recursive fusion federated learning method (LLRFed). Data is preprocessed using a sliding window and fed into a DNN with FFT. The study demonstrates that the LLRFed method significantly improves diagnosis accuracy by 9.23% accuracy in the benchmark dataset. Similarly, the fault diagnosis method for railway point machines used in high-speed and rail systems is proposed in [401], where the vibration signals from these machines under various fault conditions were analyzed using a sequential and asynchronous FL framework. The dataset included 960 vibration signal samples across sixteen conditions, each with 60 samples, and was preprocessed to ensure consistent 10-second vector lengths using trimming and zero-padding. The study employed a deep shrinkage fully convolutional network as the global model, reducing parameters by one-fourth compared to previous models, thereby minimizing communication expenses and packet loss. Data were split into training and testing sets (8:2 ratio), with one-hot encoding applied to labels. Similarly, a study by Yang et al. [402] employed an FL approach to transfer the diagnosis of industrial machines, such as bearings and robots. Their TL through distribution barycenter medium architecture integrates an FL learning framework with a server-client architecture to address data decentralization challenges, where data from different domains cannot be centrally aggregated due to privacy concerns and high transmission costs. Clients construct ResNet-based diagnosis models to represent high-level features from local data, while the server implements an StAE-based generator model to produce a distribution barycenter medium by aggregating domain-specific distribution parameters. Using balanced samples of vibration data across different health states, this method adapts marginal and conditional distributions through the generated medium samples and employs collaborative training between clients and the server to dynamically update the distribution barycenter medium. Furthermore, the study [403] presented a blockchain-based decentralized collaborative learning approach for machine fault diagnosis.1-D machinery data from two datasets, a high-speed train bogie and a shaft crack failure dataset, were used. Deep CNNs are used for the analysis in a framework that combines blockchain-based federated learning with source data-independent transition learning, in which pre-processing techniques include the use of frequency-domain information. This approach achieved over 90% testing accuracy. Moreover, the article [404] presented the use of a heterogeneous federated domain generalization network, which incorporates common representation learning to address the challenges of domain shift and privacy preservation in FD. Implementing the federated TL to achieve generalized FD across different and unseen target clients by using heterogeneous source clients, the proposed method he;ped overcome the limitations of traditional methods that rely on homogeneity among clients and the availability of target-domain data during training. The model uses a disentangled domain adaptation base model designed to minimize noise impact and enhance domain confusion, improving the extraction of fault-relevant features.

5. Challenges

In the field of MFD, despite significant academic attention, challenges remain in effectively applying advanced data-driven algorithms to real-world applications. These challenges span various aspects, including dataset issues, model architecture, and existing approaches, which are elaborated on in the subsequent subsections.

5.1. Data related challenges

Employing advanced learning approaches for MFD requires a large amount of data. However, the available datasets have many limitations that restrict the performance of the employed models. Below are the challenges researchers face when working with machinery fault datasets:

1. **Challenges related to sensors**
    - **Smart sensors:** The challenges associated with sensor networks include security and privacy, network traffic, and energy efficiency. Ensuring security and privacy is crucial, particularly in protecting data within organizational boundaries and cloud computing contexts. Network traffic is another significant challenge, as simultaneous data transmission from multiple sensors can cause congestion and potential data loss. Lastly, energy efficiency is essential for minimizing power consumption while optimizing resource utilization across the sensor network.
    - **Multi-sensor data fusion:** Challenges in multi-sensor data fusion include handling imprecision and uncertainty, integrating both homogeneous and heterogeneous sensor data, and efficiently analyzing data from diverse sources. Synchronizing data in distributed systems is critical to maintaining real-time performance. Additionally, balancing the high costs of monitoring frameworks and data acquisition systems with computational decisions between edge and cloud computing remains a key challenge.
2. **Challenges with Real-World Data:** Real-world machinery data is often noisy, inconsistent, come from heterogeneous sensors, is non-stationary, and incomplete, making fault analysis challenging. This can result from sensor failures, intermittent faults, or communication issues during data collection [237]. The contrast in data quality between real-world scenarios and controlled lab environments presents significant challenges for practical applications. Furthermore, the reluctance of industries to provide real industrial datasets for research is also an existing problem.

3. **Faulty Data Scarcity:** Real-world faulty machinery data is often scarce, making it challenging to train advanced ML models. Obtaining an adequate amount of faulty data for machinery fault detection is challenging because machinery typically operates under normal conditions. Again, the faults are of different types and severity, and obtaining enough data for each type of fault is a challenging task [1].
4. **Insufficient Labeled Data:** As previously mentioned, the majority of works rely on supervised learning, necessitating labeled data. However, precisely labeling fault types and severities is challenging, contributing to the scarcity of labeled datasets. This shortage of labeled data is a significant challenge in real-world industrial contexts, as labeling is a costly and time-consuming process, often lacking diversity to ensure effective generalization to unseen faults [25].
5. **Data Imbalance:** The faulty data scarcity leads to a class imbalance problem, with more data on normal states than failures, leading to the under-representation of certain fault types. This imbalance can hinder the learning process and bias the model towards the majority class [405].
6. **Data Incompatibility:** Detecting machinery faults becomes challenging when data comes from different sources, each with its own objectives, complexity, and criteria for data handling. Additionally, variations in data storage depth within databases create modeling issues for fault detection in machinery [406].
7. **Transferring Knowledge from Research to Practice:** Most research is conducted using publicly available datasets acquired in controlled laboratory environments. The goal is to apply this knowledge to detect unseen machinery faults, including predicting real-world faults from lab-generated data. However, several technical challenges must be addressed to achieve this [8].
8. **Problems with the available Datasets:** Most machinery fault datasets have limitations, some of which are listed below:
   - Using publicly available datasets in MFD presents various challenges, including non-classical fault recognition features, non-stationary characteristics, difficulty in accurately identifying all faults, high variance, data corruption, irregular frequency components, missing values, biases towards certain labels or classes, limited representation of fault types, restricted accessibility, and a narrow range of operating conditions [8,13,21,44].
   - Some datasets contain heterogeneous data from various sources, hindering consistency and model training. Additionally, many are complex and challenging to analyze, while some are relatively small, limiting model performance [1,34].

5.2. Challenges encountered in rotating machinery

- **Complex Movement:** Real-world machinery often involves sliding between components, which complicates the calculation of fault frequencies and affects feature informativeness.
- **Frequency Interference:** In cases where multiple types of machinery faults occur simultaneously, their interaction can lead to complex frequency interplay, making informative frequencies less clear.
- **External Noise:** Additional sources of vibration or acoustic emission or other types of noise can introduce interference and obscure relevant features.
- **Challenging Fault Types:** Some machinery faults can exhibit non-stationary behavior, lacking characteristic cyclic frequencies, making them challenging to detect.
- **Sensitivity Variability:** Features related to machinery defects can be sensitive to different operating conditions, requiring systematic adaptation.
- **Sensor Placement** Obtaining accurate machinery fault signals necessitates expensive sensors and expert involvement in securely mounting them onto machines. Inadequate data quality resulting from improper sensor placement can adversely affect the performance of DL models in fault detection [407].

5.3. Challenges in existing approaches

1. **Supervised Learning:** The predominant approach in MFD is the SL technique. However, there are certain challenges, which are mentioned as follows:
   - These methods excel in detecting known faults but struggle with new or unseen types.
   - Acquiring a diverse, well-labeled, and balanced dataset for supervised methods is resource-intensive and time-consuming, frequently resulting in limited model generalization.
   - Feature selection is crucial in supervised learning, impacting model performance. It can be challenging with high-dimensional data.
   - Supervised learning can lead to overfitting, especially when there is limited training data, causing poor performance on new data.
   - Supervised learning algorithms' sensitivity to data noise can lead to false positives and false negatives.
2. **Semi-Supervised Learning:**The use of SSL in MFD is still limited but gradually increasing. They overcome the limitations of supervised learning techniques up to some extent; however, there are certain limitations, which need to be addressed.
   - SSL needs less labeled data than supervised methods but still faces challenges in obtaining quality labels, which can be costly in MFD.
   - Challenging to determine the optimal amount of labeled data, as too little leads to poor performance, and too much can cause overfitting.
   - Combining labeled and unlabeled data consistently can introduce biases and is not always straightforward.
   - Like in supervised learning, class imbalance can be a challenge in SSL, where an unequal distribution of labeled samples across classes may bias the models.
   - Complex implementation due to the need for expertise in both supervised and unsupervised domains.
   - Sensitivity to the quality of labeled data can degrade algorithm performance.
   - Difficulty in selecting the appropriate semi-supervised learning algorithm from a variety of options.
3. **Unsupervised Learning:** The challenges in employing USL approaches in MFD are as follows:
   - The biggest advantage of unsupervised technique, i.e., not requiring labeled data, can be its limitation, as it can be difficult to evaluate the performance of the model.
   - Unsupervised techniques rely on detecting anomalies, which can be difficult with complex and diverse operational behavior.
   - Setting appropriate thresholds for anomaly detection is subjective and may require expertise.
   - These methods might generate false positives when processing complex machinery data, affecting reliability.
   - Unsupervised methods may lack clear result explanations, hindering precise fault diagnosis.

4. **Reinforcement Learning:** The limitations of RL in MFD are listed as follows:

    - Limited application of RL in MFD despite its success in other domains.
    - Underutilization of RL's potential in optimizing maintenance decisions and fault detection in existing systems.
    - Current RL algorithms for MFD often oversimplify fault diagnosis as a guessing game, resembling a classification task [408].
    - Limitations of RL algorithms in adapting to various machinery or environmental settings [117].
    - Challenges in balancing exploration and exploitation to discover optimal policies and maximize rewards in RL [365].
    - RL often requires extensive interactions with the environment, making generalization to unseen environments or tasks difficult [117].
    - Difficulties faced by RL in handling continuous action and high-dimensional state spaces.

5.4. Challenges in employing ML/DL algorithms

ML/DL algorithms have made significant strides in mapping one space, X, to another, Y, using continuous geometric transformations, particularly when a large amount of data is available. This achievement has been revolutionary across various industries [409]. Yet, human-level artificial intelligence remains an elusive goal. Challenges researchers face with deeper networks include:

1. **Feature Engineering Challenge for Classical ML:** In MFD, classical ML algorithms rely on feature engineering to identify relevant fault indicators based on operational parameters and physical characteristics. However, this approach faces challenges that can affect fault classification accuracy [8].
2. **Complex Model Training:** Training ML/DL models for machinery fault detection demands substantial computational resources and time, with deep neural networks being especially data-intensive.
3. **Model Interpretability:** Deep learning models are often regarded as **black box** models, lacking transparency and interpretability. This is a concern for safety-critical applications like machinery fault detection. Ensuring trust and facilitating decision-making in fault detection systems is crucial, necessitating research to enhance the interpretability of DL models in this context [410,411].
4. **Generalization:** Most techniques focus on specific situations rather than an integrated engineering environment, affecting their generality. Additionally, existing algorithms often struggle to find the optimal balance between approximation and generalization. Overfitting the training data can degrade performance on unseen data, while excessive generalization can lead to poor training accuracy [410,412].
5. **Data Dependency:** DL models rely heavily on large amounts of data for effective learning. They perform poorly in data-scarce situations and struggle with imbalanced datasets, often misclassifying the class with fewer samples. Moreover, DL models face challenges when learning from poor-quality or redundant data in the context of MFD [413,414].
6. **Computationally Expensive:** These models are computationally intensive and require powerful hardware for training and execution, making them costly. This can be a limiting factor for small research teams and industrial applications with limited resources [1].
7. **Feature Extraction Challenges:** Selecting relevant and important features from raw sensor data is vital for accurate fault detection, but identifying the most informative ones within a large dataset can be complex. Effective feature extraction methods and domain expertise are required to capture fault-related patterns.
8. **No Standardized Approach for Selecting DL Architectures:** Choosing the appropriate DL tool and architecture for machinery fault detection is challenging due to the absence of standardized guidelines. DL models have numerous hyperparameters, requiring expert knowledge and computational resources for optimal performance. Additionally, most existing DL architectures are designed for image data, so engineers are encouraged to explore custom convolutional architectures tailored to the unique characteristics of industrial mechanical data, which hold potential for improved fault diagnosis [415]. The limitations of some of the most commonly used DL architectures in MFD are listed below:

    - **Autoencoders:** Traditional autoencoders often include a pre-training stage to initialize network weights, and errors in the first layers can lead to the network learning to reconstruct the average of the training data. Additionally, autoencoders have limited interpretability, as the models typically learn a compressed representation.
    - **CNN:** CNNs rely on labeled data and require multiple layers to capture the entire hierarchy. They also face challenges in capturing long-term dependencies in time-series data, have fixed input size constraints that pose problems for varying-length time series, are highly sensitive to hyperparameter tuning, and struggle with irregularly sampled data or missing values.
    - **DBN:** DBNs have limitations, such as unclear optimization steps based on maximum likelihood training approximation and their inability to account for the two-dimensional structure of input images, impacting their performance and applicability in computer vision and multimedia analysis tasks [416].
    - **GAN:** GANs may encounter instability and challenges in training convergence, are susceptible to mode collapse where the generator produces only a limited variety of samples, face difficulties in evaluating the quality and performance of generated samples, and require careful tuning of both architecture and training processes [417,418].
    - **RNN and LSTM:** These models struggle with training on long sequences due to vanishing/exploding gradients, have limited ability to capture complex temporal dependencies, and present challenges in parallelizing the training process.
    - **TL:**
        - Transfer learning can help mitigate data-related challenges but may not always be suitable for MFD applications due to differences in operating conditions, sensor configurations, and machine types, which may limit the effectiveness of knowledge transfer from one task to another.
        - Improving diagnosis accuracy is challenging due to factors such as transfer errors between different working conditions, unique characteristics of different components, scale amplification issues between laboratory and industrial fields, and limitations imposed by sample size and prior knowledge.
        - Negative transfer in transfer learning can lead to poor performance in the target task. Preventing negative transfer remains an open challenge, as it is more likely to occur if the source task significantly differs from the target task. Selecting source data that closely resembles the target data can help mitigate negative transfer [21].

– Developing fault detection models applicable across different machines or equipment types is challenging. Each machine has unique operating characteristics, environmental conditions, and fault manifestations. Ensuring the generalizability of fault detection models to new, unseen machines remains a significant challenge.
– Although MFD has made significant advancements in fault recognition, feature extraction, dynamic condition monitoring, fault severity evaluation, and RUL prediction, current research indicates that deep transfer learning techniques have primarily been utilized for fault feature extraction.
– The computational burden increases due to additional complexity during transfer between source and target domains, as well as the inherent computational demands of deep learning architectures.

- **XAI:** XAI identifies important features but often fails to explain why they are critical.

5.5. Other challenges

- Lack of cross-validation and experiment repetitions in most research.
- Insufficient explanation of pseudo-code in proposed methods.
- Selecting appropriate evaluation metrics that accurately reflect the real-world impact of false positives and false negatives is challenging.
- Proper cross-validation with machinery fault data is challenging due to its temporal nature and the need to prevent data leakage.
- In some industries, regulatory requirements for implementing and validating fault detection systems add an extra layer of complexity.
- Scaling fault detection systems across an entire industrial facility with numerous machines may pose challenges in terms of hardware and infrastructure.

6. Recommendations for future researchers

6.1. Machinery and dataset enhancements

1. **Study and Analysis:** As the initial and foremost step, we recommend that future researchers conduct a thorough and detailed study of the dataset they will be working on. As mentioned multiple times, data is fundamental to the performance of any algorithm. Thus, careful analysis of the data prior to model development may lead to a better understanding of the problem and improved model performance.
2. **Dataset Creation:** When generating a dataset, careful considerations should be made.
   - Incorporate high-quality sensors like accelerometers and acoustic emission sensors for data acquisition under varying conditions (healthy and faulty states, different severities).
   - Proper sensor placement is essential for accurate data capture, requiring sensors to be positioned close to the source and securely mounted to minimize interference and noise [33].
   - Test beds should be designed to replicate real-world conditions, with carefully chosen materials and configurations.
   - Collect data across different load and speed settings to ensure meaningful fault signatures are captured.
   - Label datasets accurately with details on fault types, severities, and experimental conditions.
   - Apply preprocessing techniques such as filtering and normalization to refine data for algorithm development.
   - Organize data systematically for easy access and analysis. If sharing publicly, host datasets on reliable platforms with comprehensive documentation to support other researchers in the field.
3. **Sensor Strategy:** Adjust the number of sensors based on the chosen method. Classical methods may require fewer sensors, while deep learning approaches benefit from multiple sensors at various locations. Integrate data from diverse sources, such as vibration, temperature, acoustic signals, and torque, to enhance fault pattern recognition and model performance [32].
4. **Dataset Expansion:** Ensure adequate dataset size for training robust deep learning models. Employ data augmentation, signal processing techniques, and synthetic data generation to increase dataset diversity. Techniques like transfer learning and domain adaptation can mitigate issues related to data scarcity [406].
5. **Data Preprocessing:** Implement preprocessing techniques to remove noise, address missing data, and filter out irrelevant information. Convert raw data into frequency or time-frequency domains to reveal hidden patterns, thereby improving the efficacy of fault detection algorithms.
6. **Comprehensive Model Evaluation:** Use a mix of common benchmarks and real-world industrial data for thorough model evaluation, ensuring models are tested under diverse conditions and can handle practical challenges in machinery fault detection.
7. **Data Fusion and Handling:**
   - Enhance data collection through multi-sensor data fusion for comprehensive multi-fault diagnosis.
   - Apply cost-sensitive learning and ensemble methods with resampling strategies to manage data distribution.
   - Use datasets that reflect the imbalanced nature of real industrial settings for robust model training.

6.2. Algorithm development

- Proper assessment of the working environment and operating conditions is recommended. For simpler setups, classical ML methods or frequency-based models are suitable. In noisy or complex environments with multiple operating points, advanced DL approaches should be considered, incorporating denoising techniques for noise resilience.
- Some studies [408,419] suggest that deep learning is not always superior to traditional methods in machinery health monitoring. It is advisable to start problem-solving with simpler methods.
- Select an appropriate DL/ML architecture that balances complexity, interpretability, and computational requirements for a specific machinery fault detection task [24].
- Unsupervised anomaly detection methods can address the scarcity of labeled data and class imbalance problems. Since faulty data exhibit anomalous patterns, anomaly detection methods can identify them without requiring labeled data. Additionally, they may detect new and unseen faults.
- Semi-supervised or unsupervised learning algorithms can also mitigate labeled data scarcity. When using semi-supervised learning, employ an appropriate proportion of labeled data while leveraging unlabeled instances.
- Besides data augmentation techniques, the data imbalance problem can be addressed through the design of loss functions. For instance, the Pareto optimization-based approach in [420] introduced a multi-task learning framework to balance trade-offs between tasks, effectively handling data imbalance. Similarly, adaptive cross-entropy loss automatically reweights the influence of majority and minority samples during training, improving model performance on imbalanced data.

- Apply regularization techniques and denoising methods to clean signals, avoid overfitting, and improve the generalization performance of DL models in machinery fault detection.
- Conduct comprehensive testing using cross-validation methods and repeated experiments to validate model performance, stability, and generalizability.
- Integrate DL with traditional ML, signal processing, feature engineering, and other techniques to leverage their combined strengths for a robust machinery fault detection algorithm [167].
- Investigate domain adaptation methods to improve the effectiveness of transfer learning in machinery fault detection applications, considering different operating conditions and machine configurations.
- For reinforcement learning applications, explore the development of new reward functions, algorithms, and learning paradigms suited to machinery fault detection challenges.
- Incorporate explainable AI techniques to enhance the interpretability of deep learning models, providing insights into their decision-making processes and increasing confidence in their predictions.

6.3. Other recommendations

- Expand research to cover a wider range of faults and machinery. While existing studies predominantly focus on specific machine components like bearings, gears, and motors, real-world applications frequently involve these components operating together. This necessitates the development of diagnostic methods that can assess interactions between various machine parts in combined systems.
- To enhance the accessibility and comparability of methods, we strongly recommend publishing the source code of machinery fault detection algorithms. Open-source code allows other researchers to implement, evaluate, and build upon existing work, promoting transparency, collaboration, and innovation in the field.
- Researchers should adopt a thorough and structured documentation approach to benefit future researchers in the field.
- Consider incorporating online learning techniques in machinery fault detection to enable continuous adaptation to changing machine conditions and new failure types [117].
- Employing digital twins can help simulate scenarios to understand root causes of failures, predict faults for proactive maintenance, and enable real-time monitoring [319].
- Although open-source codes in this field are scarce, a few researchers maintain open-source resources. We recommend that future researchers visit these sites for assistance: https://qinyi-team.github.io, https://biswajitsahoo1111.github.io/, https://github.com/liguge,and so on. We also urge future researchers to make their code and datasets open-source to benefit the entire research community.

7. Future prospects

The future of machinery fault diagnosis holds several promising opportunities, with developments aimed at improving the accuracy, dependability, and adaptability of diagnostic systems in industrial settings. Researchers are focusing on enhancing the robustness of fault detection and classification by developing new methods to detect crack damage under mixed-speed conditions. This work highlights the importance of noise-resilient algorithms, which are crucial for accurate diagnosis in noisy industrial environments. Furthermore, the integration of advanced models such as transformer architectures and physics-informed neural networks is set to advance MFD systems. Researchers are also exploring various data fusion methods–including sensor, feature, and decision fusion–to merge different data types such as vibration, current, torque, acoustic signals, and visual images. This approach optimizes the preprocessing of these data types, improving the accuracy and efficiency of fault detection. Thus, the development of methods to handle complexity across different fusion levels is a prospective area for future research.

The exploration of semi-supervised and reinforcement learning in MFD is emerging, focusing on reducing dependence on labeled data and employing sequence learning to predict faults and enhance detection efficiency. Future research in MFD is increasingly emphasizing the practical applications of offline RL, aiming for better generalization in new environments. This shift prioritizes detecting faults over simply classifying them, reflecting a more dynamic and proactive approach to fault management.

Furthermore, the integration of advanced techniques such as domain generalization, domain confusion, and domain adaptation is set to address unforeseen machinery fault conditions, data imbalance, and limited labeled data. Additionally, the development of XAI approaches promises to make MFD more transparent and interpretable, building trust in the decision-making processes of these systems. However, while XAI identifies which features are important, it often does not explain why those features are critical, highlighting a need for further research to bridge this gap and provide deeper insights into feature significance. Overall, the future of MFD research is characterized by innovation and practicality, promising more reliable, accurate, and adaptive fault detection methods for industrial applications.

Last but not least, to tackle the computational challenges of real-time diagnostics, refining data handling capabilities and computational efficiency is crucial. Methods such as pre-training, neural architecture search, and model compression will play a key role. Integrating mechanistic knowledge with deep learning is expected to improve both model interpretability and generalization, leading to more autonomous and sophisticated monitoring systems. The incorporation of digital twin technology and hybrid data-driven approaches will also play a significant role in minimizing prediction errors and refining maintenance strategies. Collectively, these future research possibilities will advance MFD not only to meet the complex demands of modern industries but also to progress toward more reliable and efficient fault detection methodologies.

8. Conclusion

This review provides a comprehensive overview of machinery fault diagnosis approaches, covering topics from data sources to predictive maintenance strategies. It explores a wide range of datasets in the field and discusses both traditional and advanced methodologies, including deep learning, federated learning, reinforcement learning, transfer learning, and physics-informed neural networks. The review highlights the critical role of data collection and maintenance analysis, emphasizing the potential of various predictive methods to enhance fault detection and machine reliability. The identified challenges serve as key areas for future research, while the recommendations provided aim to improve dataset quality, algorithm selection, and practical applications.

This review serves as a valuable resource for both newcomers and experts in the field of machinery fault detection, contributing to advancements in fault detection and machine reliability across various industries.

CRediT authorship contribution statement

Dhiraj Neupane: Writing – review & editing, Writing – original draft, Methodology, Data curation, Conceptualization. **Mohamed Reda Bouadjenek:** Supervision. **Richard Dazeley:** Supervision, Conceptualization. **Sunil Aryal:** Supervision, Methodology, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Abbreviations used in the article and their full forms

Table A.13
Abbreviations and their Full Forms.

Abbreviation	Full form	Abbreviation	Full form
AcE	Acoustic Emission	ACGANL	Auxiliary Classifier Generative Adversarial Network
AD	Anomaly Detection	AE	Autoencoder
ANN	Artificial Neural Network	AT	Automobile Transmission
AUC	Area Under the Curve	BEGAN	Boundary Equilibrium Generative Adversarial Network
BiRNN	Bidirectional Recurrent Neural Network	BiGRU	Bidirectional Gated Recurrent Units
BPNN	Backpropagation Neural Network	CAE	Contractive AE
Cap-Net	Capsule Network	CbM	Condition-based Maintenance
CDCN	Coupled Dense Convolutional Network	C-MAPSS	Commercial Modular Aero-Propulsion System Simulation
CNN	Convolutional Neural Network	CNNFD	CNN-based Fault Diagnosis
CGAN	Conditional Generative Adversarial Network	ConvAE	Convolutional AE
CPF	Chemical Process Faults	CRNN	Convolution Recurrent Neural Network
CS	Coarse-grained Similarity	CWRU	Case Reserve Western University
CWT	Continuous Wavelet Transform	DAE	Denoising AE
DBN	Deep Belief Network	DCNN	Deep CNN
DCGAN	Deep Convolutional Generative Adversarial Network	DFT	Discrete Fourier Transform
DIRG	Dynamic and Identification Research Group	DL	Deep Learning
DNN	Deep Neural Network	DQN	Deep Q-Networks
DRL	Deep Reinforcement Learning	DT	Decision Tree
DSTG	Defence Science and Technology Group	EBGAN	Energy-Based Generative Adversarial Network
EMD	Empirical Mode Decomposition	ER	Experience Replay
FD	Fault Diagnosis	FEMTO-ST	Franche-Comté Électronique Mécanique, Thermique et Optique - Sciences et Technologies
FFT	Fast Fourier Transforms	FGDAE	Full Graph Dynamic AE
FGPAA	Fault-Attention Generative Probabilistic Adversarial AE	FTD-MSDA	Feature-level and Task-specific Distribution Alignment Multisource Domain Adaptation
GAN	Generative Adversarial Network	GRU	Gated Recurrent Units
HHT	Hilbert-Huang Transform	HMM	Hidden Markov Model
HTSR	Health-Adaptive Time-Scale Representation	HUMS	Health Usage Monitoring System
iF	Isolation Forest	IMS	Intelligent Maintenance Systems
IRT	Infrared Thermal	JNU	Jiangnan University
kNN	K-Nearest Neighbors	LOF	Local Outlier Factor
LS-GAN	Least Squares GAN	LS-SVM	Least Square Support Vector Machine
LSTM	Long Short-Term Memory	MaFaulDa	Machinery Fault Database
MCSA	Motor Current Signature Analysis	MFD	Machinery Fault Diagnosis
MFPT	Motor Fault Prognostic Testbed	MFS	Machinery Fault Simulator
MFS-RDS	MFS- Rotor Dynamics Simulator	MHM	Machine Health Monitoring
ML	Machine Learning	MLP	Multilayer Perceptron
ML-TRL	Multi-Label Transfer Reinforcement Learning	MMD	Mean Maximum Discrepancy
MWK	Multi-Wavelet Kernel	NAS	Neural Architecture Search
NBC	Naive Bayes Classifier	NEU	North East University
NREL	National Renewable Energy Laboratory	OC-SVM	One Class Support Vector Machines
PCA	Principal Component Analysis	PdM	Predictive Maintenance
PHM	Prognostics and Health Management	PINN	Physics Informed Neural Networks
PM	Physical Model	PnM	Preventive Maintenance
PPO	Proximal Policy Optimization	PSO	Particle Swarm Optimization
PU	Paderborn University	RADIS	Real-time Anomaly Detection Intelligent System
RBM	Restricted Boltzmann Machines	RF	Random Forest
RL	Reinforcement Learning	RM	Reactive Maintenance
RNN	Recurrent Neural Network	RPM	Revolutions per Minute
RUL	Remaining Useful Life	SCADA	Supervisory Control and Data Acquisition
SDU	Shandong University	SEU	Southeast University
SL	Supervised Learning	SpAE	Sparse AE
SQ	SpectraQuest	SRU	Simple Recurrent Units
SSL	Semi-Supervised Learning	StAE	Stacked AE
STFT	Short-Time Fourier Transform	SVM	Support Vector Machines
TEP	Tennessee Eastman Process	THU	Tsinghua University
TL	Transfer Learning	TML	Traditional Machine Learning
UA-FS	University of Alberta - Fixed Shaft	UESTC	University of Electronic Science and Technology of China
UO	University of Ottawa	UoC	University of Connecticut
UORED-VAFCLS	University of Ottawa Rolling-element Dataset – Vibration and Acoustic Faults under Constant Load and Speed Conditions	USL	Unsupervised Learning
VAE	Variational AE	WGAN	Wasserstein GAN
WK-CNN	Wide-Kernel CNN	WPT	Wavelet Packet Transform
WSCN	Weight-Shared Capsule Network	WT	Wavelet Transforms
XAI	Explainable Artificial Intelligence	XJTU-SY	Xi’an Jiaotong University and Changxing Sumyoung Technology

Appendix B. Machine fault detection datasets

Airbus helicopter accelerometer dataset

Available at ETHZurichRepository, this dataset includes vibration data from Airbus helicopters, with 1-minute sequences recorded at 1024 Hz. It comprises 1677 training sequences and 594 validation sequences, with half labeled as abnormal. The dataset supports automated flight test data validation and was used in Airbus's 2019 AI challenge to classify sequences as healthy or faulty [34].

AI4I 2020 Predictive Maintenance Dataset

The AI4I 2020 Predictive Maintenance Dataset is a synthetic dataset designed for predictive maintenance in the computer industry. It contains 10,000 data points with 14 features each, such as product ID, air and process temperatures, rotational speed, torque, tool wear, and a 'machine failure' label. The label identifies if a failure has occurred, without specifying the failure mode—tool wear, heat dissipation, power, overstrain, or random. For dataset access, visit thislink [35, 36].

C-MAPSS dataset

The C-MAPSS dataset, derived from the Commercial Modular Aero-Propulsion System Simulation (C-MAPSS), is pivotal for developing predictive models in aerospace fault detection. This dataset is sourced from a dynamic simulator designed for large commercial turbofan engines and includes 21 sensor variables like temperature, pressure, and speed across five subsets depicting various wear levels and degradation processes. It supports testing of advanced algorithms through user-defined transient simulations and linear state-space models. Although ideal for training deep learning models due to its extensive training samples and diverse operating conditions, it is important to note that as a simulated dataset, it may not fully replicate real-world scenarios. The dataset is accessible for download at NASADataPortal.

CWRU bearing dataset:

The CWRU bearing dataset, widely utilized for MFD research, comprises experimental data featuring faults like inner race, outer race, and ball faults with varying severities under different load conditions. This dataset provides time-domain vibration signals sampled at two different frequencies – 12 KHz and 48 KHz – and collected by sensors at different positions. Data encompasses single-point faults introduced via electro-discharge machining, with fault diameters ranging from 7 to 40 mils, motor loads from 0 to 3 horsepower, and speeds between 1720 to 1797 rpm. Available publicly at CaseWesternReserveUniversity'ssite, files are named by fault type, size, and load, such as `B007_0` for a 7 mil ball fault under no load. This comprehensive dataset is crucial for validating fault diagnosis algorithms under various conditions [44].

DIRG bearing dataset

The Dynamic and Identification Research Group (DIRG) at Politecnico di Torino's Department of Mechanical and Aerospace Engineering provides a dataset on high-speed aeronautical roller bearings tested above 6000 rpm with two accelerometers positioned at points A1 and A2. This dataset includes two experimental sessions: the first tests various damages on bearing B1, such as localized faults on the inner race ring or a single roller, under varied speeds and loads; the second involves a prolonged test of a single damaged bearing over approximately 330 h at constant speed and load, recording vibration signals at a sampling rate of 51,200 Hz. More details and the dataset download are available in [22,100] at thislink.

EDGFD

The Experimental Dataset for Gear Fault Diagnosis (EDGFD) dataset includes radial vibration signals from a gearbox with helical gears under three conditions: healthy, one chipped tooth, and three worn teeth. It features a helical gear system composed of a 15-teeth pinion and a 110-teeth wheel, operating at a speed ratio of 7.33 and a nominal pinion speed of 1420 RPM. The gear mesh frequency typically measures at 355 Hz but was recorded at 365 Hz under test conditions. Data was captured over a 10-second duration in each test scenario and saved in MATLAB mat-file format. The dataset is publicly accessible at thislink. More details are available in the article "Experimental Dataset for Gear Fault Diagnosis" [102].

FEMTO bearing dataset

The FEMTO bearing dataset, also called as **IEEE 2012 Data Challenge Dataset**, provided by the FEMTO-ST Institute in France, comprises time-domain vibration signals, temperature measurements, and speed data from bearings under both normal and accelerated degradation conditions, collected on the PRONOSTIA experimental platform. This platform is designed to accelerate bearing degradation under controlled conditions for real-time monitoring, using a rotational speed sensor and a force sensor to characterize operating conditions. Vibration signals are recorded every 10 s at 25.6 kHz, while temperature is logged at 10 Hz. The dataset includes 17 failure instances under three different conditions, essential for developing and validating prognostic models for bearing RUL prediction [30,116]. The dataset is available at thislink.

Gearbox Fault Diagnosis Data

The gearbox fault diagnosis dataset comprises vibration data collected using SpectraQuest's Gearbox Fault Diagnostics Simulator. This dataset was obtained using four vibration sensors placed in four distinct directions, capturing data under various load conditions ranging from 0% to 90%. The dataset consists of two different scenarios: (1) Healthy condition and (2) Broken Tooth Condition. Ten separate text files are available for each case, resulting in a comprehensive collection of data for gearbox fault diagnosis [421]. The information related to dataset can be found at thiswebsite. And, the dataset can be downloaded from *data.world*, particularly from thiswebsite.

HUMS dataset

The Defence Science and Technology Group (DSTG), Melbourne, Australia also has been conducting data challanges on Health and Usage Monitoring System (HUMS). In 2023, DSTG organized HUMS data challaenge 2023, and the dataset contains vibration data from a planet gear fatigue crack propagation test on a Bell Kiowa 206B-1 helicopter gearbox. The test was conducted at DSTG's Melbourne facility in January 2022, simulating fatigue cracking in helicopter gears at speeds up to 6,000 RPM. The dataset is accessible at HUMSDataChallenge2023 [422]. Similarly, they have again announced for the data challenge for 2025. This dataset was generated from the Bell 206B-1 Kiowa (OH-58) helicopter main rotor gearbox test program in their helicopter transmission test facility. The link to download the dataset is HUMSDataChallenge2025.

IMS bearing dataset:

The IMS Bearing dataset from the University of Cincinnati's Center for Intelligent Maintenance Systems is crucial for bearing prognostics. It includes high-resolution time-domain vibration signals and temperature data from run-to-failure tests, captured by accelerometers mounted on test bearings. Data files record one-second vibration snapshots at 20 kHz, facilitating the development and evaluation of data-driven prognostic models for machinery fault detection and prediction [22]. The dataset is accessible from NASA's prognostic data repository or via Kaggle at thislink.

IMTI dataset

The Induciton Motor Thermal Image (IMTI) dataset was developed by Najafi et al. [113], featuring 369 thermal images of three-phase induction motors under 11 different fault conditions, captured with a Dali-tech T4/T8 infrared thermal imager at a resolution of 320 × 240. The dataset, which includes faults like rotor blockage, cooling fan failure, and various stator winding short-circuits, aims to aid in condition monitoring of electrical equipment.

JNU bearing dataset

Jiangnan University (JNU) provided comprehensive bearing datasets for analysis and fault detection. The datasets consist of three vibration datasets, containing one health state and three fault modes which include inner ring fault, outer ring fault, and rolling element fault, recorded at a sampling frequency of 50 kHz, encompassing various rotating speeds. Therefore, the total number of classes was equal to twelve according to different working conditions [116].

Mafaulda

The *MaFaulDa* dataset comprises 1951 multivariate time series collected from a SpectraQuest machinery fault simulator. The dataset represents six simulated states: normal function, imbalance fault, horizontal and vertical misalignment faults, and inner and outer bearing faults. Data were gathered using accelerometers, a tachometer, and a microphone, monitoring two bearing-supported shafts. Faults such as rotor imbalances, shaft misalignments, and defective bearings were systematically introduced. The dataset is available at SMTUFRJlink and Kagglelink.

MFPT bearing dataset

The Motor Fault Pronostic Testbed (MFPT) bearing dataset, provided by the Society for Machinery Failure Prevention Technology, features data from a test rig using a NICE bearing (roller diameter: 0.235, pitch diameter: 1.245, number of elements: 8, contact angle: 0). This dataset captures various fault conditions and baseline measurements, including three baseline conditions, ten fault conditions across outer and inner races under varying loads, and specific data on real-world examples from wind turbine bearings. Data are available *.mat* format and include load, shaft rate, and sample rate details. Accompanied by sample code, this dataset aims to support researchers and CBM practitioners in advancing their techniques. Further information and downloads are available at MFPTwebsite.

NEU dataset

North East University dataset comprises 1,800 steel surface defect images, encompassing six distinct fault types, including crazing, patches, rolled-in scale, inclusion, pitted surface, and scratches. Notably, the dataset maintains a balanced distribution among these fault types, making it a valuable resource for training and testing fault detection models [423]. This dataset can be downloaded from thiswebsite.

NREL wind turbine gearbox condition monitoring vibration analysis benchmarking datasets

The NREL Wind Turbine Gearbox Condition Monitoring Vibration Analysis Benchmarking Datasets by the National Renewable Energy Laboratory (NREL) address the scarcity of benchmarking datasets for wind turbine CBM systems. This initiative, part of the Gearbox Reliability Collaborative, includes data from both a healthy and a similarly designed damaged gearbox that experienced loss-of-oil events, affecting its bearings and gears [120]. Data from these gearboxes, which underwent dynamometer testing, are available for download here. The test gearboxes, capable of operating at 1800 rpm and 1200 rpm, were outfitted with over 125 sensors, providing extensive data to support research and development in vibration-based condition monitoring techniques. This dataset is crucial for validating and advancing CBM systems in the wind industry.

PHM2009 gearbox dataset

The PHM-2009 gearbox dataset, shared by the IEEE international conference on the PHM 2009 data challenge, is valuable for evaluating gearbox fault detection algorithms. This dataset contains 20 test cases with vibration signals, temperature and torque measurements. Preprocessing and feature extraction are required before the dataset can be used for machine learning algorithms. Researchers have developed fault detection algorithms using various traditional and advanced methods and compared their performance using evaluation metrics. The dataset remains an important benchmark for further development of error detection methods [424]. This dataset can be downloaded from thislink.

PHM2010 milling cutter dataset

The dataset from the 2010 PHM Society Conference Data Challenge focuses on the RUL estimation for CNC milling machine cutters. Utilizing a Kistler quartz 3-component dynamometer and three accelerometers, it captures cutting forces and vibrations across the X, Y, Z axes, while an acoustic emission sensor monitors high-frequency stress waves during cutting. Operational conditions include a spindle speed of 10400 rpm, a feed rate of 1555 mm/min in the x-direction, and cutting depths of 0.125 mm and 0.2 mm in the y and z directions, respectively, with data sampled at 50 kHz [425]. Despite its collection under a dry milling setup which deviates from typical industrial conditions, the dataset offers valuable insights for RUL prediction in milling operations. It includes 3D cutting force and vibration data along with acoustic signals, supporting both single-sensor and multi-sensor fusion analyses. However, only three of the six cutter datasets are labeled, and the uniform milling condition limits the diversity for cross-prediction scenarios.

PHME datasets

The European conference on Prognostics and Heath Managements (PHME) has been conducting data challenges. In the year 2021, the challenge focused on fault detection, classification, and root cause identification in a manufacturing production line setup. Participants were provided with real-world industrial datasets to apply state-of-the-art algorithms and models. Similarly, participants were tasked with solving a classification problem related to a real production line equipped with Industry 4.0 technologies in the year 2022. The extensive dataset was provided by Bitron Spa, focusing on the operation of automated and integrated machines. Furthermore, the recent data challenge in the year 2024 involved estimating the RUL of a subway ticket validation door system. Participants were encouraged to develop models to predict system failure based on the degradation of door position over time without knowing the operating conditions. All the dataset can be found in their GitHubPage.

PU bearing dataset

The PU (Paderborn University) bearing dataset, provided by the KAT data center in Paderborn University includes high-resolution vibration data, torque, and temperature measurements from a custom-built test rig that simulates various bearing faults under different conditions. It comprises data from six healthy bearings and 26 damaged sets, with damage induced both artificially and through accelerated life tests, and is designed to enhance ML algorithm testing. The complete dataset is available for download here.

Rotor fault dataset

The dataset was acquired using a rotor lab bench setup consisting of a DC motor connected to an 850 mm rotor via bearings and couplings, with two mass disks (75 mm diameter) and a GTS3-TG series simulator for data acquisition. Data capture involved two eddy current sensors processing signals for amplification and filtering by a processor before storage on a computer. Four rotor states—normal, unbalanced, misaligned, and rubbing—were recorded at 1200 r/min with a sampling frequency of 2048 Hz and 1-second sample length. Unbalance was simulated with a 2 g mass on the disk; misalignment by adjusting the shaft coupling; rubbing by introducing a screw contact with the shaft. The dataset comprises 180 samples across 45 test groups, split into 80 training and 100 testing samples, each processed via wavelet thresholding and stored as a 2-D matrix. The data is available at MendeleyData.

SEU gear fault dataset

The SEU gear fault dataset from Southeast University, China, is collected using the Drivetrain Dynamics Simulator. It consists of two subdatasets: bearing and gear data under conditions 20-0 and 30-2 for speed-load. It features motor and gearbox vibrations (x, y, z), motor torque, and includes five gear fault types (healthy, chipped, missing tooth, root, surface) and five bearing faults (healthy, inner, outer, combined, rolling element). This dataset is crucial for studying time-frequency distributions and fault diagnosis. Available for download at GitHub [73].

TEP simulation dataset

The Tennessee Eastman Process (TEP) is a chemical simulation benchmark used for fault diagnosis in continuous processes, involving 22 simulation runs with 52 variables, and 20 fault types. Available at IEEEDataport.

THU gearbox dataset

The THU dataset from Tsinghua University comprises vibration data from gearbox gear fault experiments conducted in 2019 using an HS-200 single-stage planetary gearbox. The setup included a motor (29-31 Hz), a planetary gearbox with artificially damaged gears, and two accelerometers mounted on the gearbox case in the x and y directions. It features nine fault types, including a healthy condition and various levels of damage to sun and planetary gears. While the data provides a broad spectrum for fault analysis, it is derived from a controlled environment, which may not fully replicate real-world conditions [41,136].

UA-fs gearbox dataset

Generated by the University of Alberta (UA), this dataset contains vibration and speed signals from a fixed-shaft (FS) gearbox, highlighting five levels of tooth root crack severity. Data were collected under two speed profiles with a total of 255,037 samples.

UO bearing dataset

Provided by the University of Ottawa (UO), this dataset features acceleration and speed signals from bearings under various health conditions. It is available at MendeleyData.

UoC gearbox fault dataset

This dataset is provided by University of Connecticut (UoC). The dataset contains a collection of vibration data from two-stage gears under various operating conditions, and includes data from healthy conditions as well as conditions with various gear defects such as missing teeth, root cracks, spalling and flaking chips. The dataset can be downloaded from the UoC website. The sampling frequency was 20 kHz [41,426,427]. The UOC gear failure dataset is a valuable resource for researchers and technicians involved in gear failure diagnosis. It can be used to train and evaluate machine learning models for transmission fault detection. The dataset can also be used to develop new methods for gearbox fault diagnosis.

UORED-VAFCLS dataset

University of Ottawa Rolling-element Dataset – Vibration and Acoustic Faults under Constant Load and Speed conditions (UoARED-VAFCLS) dataset offers vibration and acoustic data for fault analysis under constant load and speed, available at MendeleyData.

Wind turbine SCADA dataset

The dataset contains supervisory control and data acquisition (SCADA) data collected at wind farms. SCADA systems record various measurements, including wind speed, power output and rotor speed. The dataset helps to analyze wind turbine performance and detect abnormal behavior, such as gearbox, generator or other component failures [146]. Some of the SCADA datasets are:

(i) Operation SCADA dataset of an urban small wind turbine in S∼ ao Paulo, Brazil
(ii) 2018ScadaDataofaWindTurbineinTurkey (Kaggle)
(iii) WindTurbineSCADAopendata (GitHub).

XJTU-SY bearing dataset

The XJTU-SY bearing dataset is provided by Xi'an Jiaotong University and Changxing Sumyoung Technology Co. It includes run-to-failure data from 15 rolling element bearings under three conditions: 2100 rpm with 12 kN, 2250 rpm with 11 kN, and 2400 rpm with 10 kN. Tests were conducted using an AC motor, a speed controller, a shaft, support bearings, and a hydraulic loader. Data from two accelerometers at 25.6 kHz, capturing 32768 points per minute, are stored in CSV files. This dataset is ideal for validating prognostics algorithms and can be accessed at thislink.

MFS

Machine fault simulators (MFS) are critical tools for studying machinery faults in controlled environments, used extensively for over two decades. These simulators integrate robust multi-channel data acquisition systems and comprehensive data analysis software, supporting a variety of analyses such as Time Waveform, Amplitude Spectrum, and Frequency Response [130]. SpectraQuest (SQ) and TIERA are leading providers, offering models like the MFS–Lite, Machinery Fault & Rotor Dynamics Simulator, and MFS–Magnum. SQ's systems are noted for their versatility and ease of use, while TIERA's feature modular designs that facilitate the study of common faults like unbalance and bearing defects without halting production. These tools are invaluable for advancing practical knowledge in machinery diagnosis and fault detection, with datasets like MaFaulDa enhancing research capabilities.

Noteworthy mentions

Several other datasets have been utilized in MFD, including the HOUDE Dataset [72], Ball Screw Dataset [72], SDU Dataset [78], UNSW Turbine Blade Dataset [91], Qianpeng Company Gearbox Dataset [126], NCEPU Gear Dataset [96], QPZZ-II Dataset [385], and the Locomotive Bearing Dataset [74]. For a comprehensive list of relevant datasets used in MFD, you can visit this GitHubPage.

Data availability

No data was used for the research described in the article.

References

[1] Dhiraj Neupane, Jongwon Seok, Bearing fault detection and diagnosis using case western reserve university dataset with deep learning approaches: A review, IEEE Access 8 (2020) 93155–93178.
[2] Chuan Li, Shaohui Zhang, Yi Qin, Edgar Estupinan, A systematic review of deep transfer learning for machinery fault diagnosis, Neurocomputing 407 (2020) 121–135.
[3] Hongtian Chen, Hao Luo, Biao Huang, Bin Jiang, Okyay Kaynak, Transfer learning-motivated intelligent fault diagnosis designs: A survey, insights, and perspectives, IEEE Trans. Neural Netw. Learn. Syst. (2023).
[4] Carlos Ferreira, Gil Gonçalves, Remaining useful life prediction and challenges: A literature review on the use of machine learning methods, J. Manuf. Syst. 63 (2022) 550–562.
[5] Oguzhan Das, Duygu Bagci Das, Derya Birant, Machine learning for fault analysis in rotating machinery: A comprehensive review, Heliyon (2023).
[6] Paulina Gackowiec, General overview of maintenance strategies–concepts and approaches, Multidiscip. Asp. Prod. Eng. 2 (1) (2019) 126–139.
[7] Yaguo Lei, Bin Yang, Xinwei Jiang, Feng Jia, Naipeng Li, Asoke K. Nandi, Applications of machine learning to machine fault diagnosis: A review and roadmap, Mech. Syst. Signal Process. 138 (2020) 106587.
[8] Shen Zhang, Shibo Zhang, Bingnan Wang, Thomas G. Habetler, Deep learning algorithms for bearing fault diagnostics—A comprehensive review, IEEE Access 8 (2020) 29857–29881.
[9] Duy-Tang Hoang, Hee-Jun Kang, A survey on deep learning based bearing fault diagnosis, Neurocomputing 335 (2019) 327–335.
[10] Shiza Mushtaq, M.M. Manjurul Islam, Muhammad Sohaib, Deep learning aided data-driven fault diagnosis of rotatory machine: A comprehensive review, Energies 14 (16) (2021) 5150.
[11] Jaskaran Singh, Moslem Azamfar, Fei Li, Jay Lee, A systematic review of machine learning algorithms for prognostics and health management of rolling element bearings: fundamentals, concepts and applications, Meas. Sci. Technol. 32 (1) (2020) 012001.
[12] Shengnan Tang, Shouqi Yuan, Yong Zhu, Deep learning-based intelligent fault diagnosis methods towards rotating machinery, IEEE Access 8 (2020) 9335–9346.
[13] Cem Ekin Sunal, Vladimir Dyo, Vladan Velisavljevic, Review of machine learning based fault detection for centrifugal pump induction motors, IEEE Access 10 (2022) 71344–71355.
[14] Prashant Kumar, Ananda Shankar Hati, Review on machine learning algorithm based fault detection in induction motors, Arch. Comput. Methods Eng. 28 (2021) 1929–1940.
[15] Omar AlShorman, Muhammad Irfan, Nordin Saad, D. Zhen, Noman Haider, Adam Glowacz, Ahmad AlShorman, A review of artificial intelligence methods for condition monitoring and fault diagnosis of rolling element bearings for induction motor, Shock. Vib. 2020 (2020) 1–20.
[16] Georg Helbing, Matthias Ritter, Deep learning for fault detection in wind turbines, Renew. Sustain. Energy Rev. 98 (2018) 189–198.
[17] Adrian Stetco, Fateme Dinmohammadi, Xingyu Zhao, Valentin Robu, David Flynn, Mike Barnes, John Keane, Goran Nenadic, Machine learning methods for wind turbine condition monitoring: A review, Renew. Energy 133 (2019) 620–635.
[18] Dileep Kumar Soother, Jawaid Daudpoto, Nicholas R Harris, Majid Hussain, Sanaullah Mehran, Imtiaz Hussain Kalwar, Tanweer Hussain, Tayab Din Memon, The importance of feature processing in deep-learning-based condition monitoring of motors, Math. Probl. Eng. 2021 (2021) 1–23.
[19] Shengnan Tang, Shouqi Yuan, Yong Zhu, Data preprocessing techniques in convolutional neural network based on fault diagnosis towards rotating machinery, IEEE Access 8 (2020) 149487–149496.
[20] Tianci Zhang, Jinglong Chen, Fudong Li, Kaiyu Zhang, Haixin Lv, Shuilong He, Enyong Xu, Intelligent fault diagnosis of machines with small & imbalanced data: A state-of-the-art review and possible extensions, ISA Trans. 119 (2022) 152–171.
[21] Weihua Li, Ruyi Huang, Jipu Li, Yixiao Liao, Zhuyun Chen, Guolin He, Ruqiang Yan, Konstantinos Gryllias, A perspective survey on deep transfer learning for fault diagnosis in industrial scenarios: Theories, applications and challenges, Mech. Syst. Signal Process. 167 (2022) 108487.
[22] Huailiang Zheng, Rixin Wang, Yuantao Yang, Jiancheng Yin, Yongbo Li, Yuqing Li, Minqiang Xu, Cross-domain fault diagnosis using knowledge transfer strategy: A review, IEEE Access 7 (2019) 129260–129290.
[23] Guokai Liu, Weiming Shen, Liang Gao, Andrew Kusiak, Knowledge transfer in fault diagnosis of rotary machines, IET Collab. Intell. Manuf. 4 (1) (2022) 17–34.
[24] Mohammed Hakim, Abdoulhdi A. Borhana Omran, Ali Najah Ahmed, Muhannad Al-Waily, Abdallah Abdellatif, A systematic review of rolling bearing fault diagnoses based on deep learning and transfer learning: Taxonomy, overview, application, open challenges, weaknesses and recommendations, Ain Shams Eng. J. 14 (4) (2023) 101945.
[25] Tim Salimans, Ian Goodfellow, Wojciech Zaremba, Vicki Cheung, Alec Radford, Xi Chen, Xi Chen, Improved techniques for training GANs, in: D. Lee, M. Sugiyama, U. Luxburg, I. Guyon, R. Garnett (Eds.), in: Advances in Neural Information Processing Systems, vol. 29, Curran Associates, Inc., 2016.
[26] Tongyang Pan, Jinglong Chen, Tianci Zhang, Shen Liu, Shuilong He, Haixin Lv, Generative adversarial network in mechanical fault diagnosis under small sample: A systematic review on applications and future perspectives, ISA Trans. 128 (2022) 1–10.
[27] Masoud Jalayer, Carlotta Orsenigo, Carlo Vercellis, Fault detection and diagnosis for rotating machinery: A model based on convolutional LSTM, fast Fourier and continuous wavelet transforms, Comput. Ind. 125 (2021) 103378.
[28] Zheng Yang, Binbin Xu, Wei Luo, Fei Chen, Autoencoder-based representation learning and its application in intelligent fault diagnosis: A review, Measurement 189 (2022) 110460.
[29] Jinchuan Qian, Zhihuan Song, Yuan Yao, Zheren Zhu, Xinmin Zhang, A review on autoencoder based representation learning for fault detection and diagnosis in industrial processes, Chemometr. Intell. Lab. Syst. (2022) 104711.
[30] Jinyang Jiao, Ming Zhao, Jing Lin, Kaixuan Liang, A comprehensive review on convolutional neural network in machine fault diagnosis, Neurocomputing 417 (2020) 36–63.
[31] Ruonan Liu, Boyuan Yang, Enrico Zio, Xuefeng Chen, Artificial intelligence for fault diagnosis of rotating machinery: A review, Mech. Syst. Signal Process. 108 (2018) 33–47.
[32] Zhiqin Zhu, Yangbo Lei, Guanqiu Qi, Yi Chai, Neal Mazur, Yiyao An, Xinghua Huang, A review of the application of deep learning in intelligent fault diagnosis of rotating machinery, Measurement 206 (2023) 112346.
[33] Amiya Ranjan Mohanty, Machinery condition monitoring principles and practices, first ed., CRC Press, Boca Raton, FL, 2014.
[34] Kilian Vos, Zhongxiao Peng, Christopher Jenkins, Md Rifat Shahriar, Pietro Borghesani, Wenyi Wang, Vibration-based anomaly detection using LSTM/SVM approaches, Mech. Syst. Signal Process. 169 (2022) 108752.
[35] S. Matzka, AI4i 2020 predictive maintenance dataset, UCI Mach. Learn. Repos. (2020).
[36] Andrea Torcianti, Stephan Matzka, Explainable artificial intelligence for predictive maintenance applications using a local surrogate model, in: 2021 4th International Conference on Artificial Intelligence for Industries (AI4I), IEEE, 2021, pp. 86–88.
[37] Wennian Yu, I.I. Yong Kim, Chris Mechefske, Remaining useful life estimation using a bidirectional recurrent neural network based autoencoder scheme, Mech. Syst. Signal Process. 129 (2019) 764–780.
[38] Wennian Yu, Il Yong Kim, Chris Mechefske, Analysis of different RNN autoencoder variants for time series classification and machine prognostics, Mech. Syst. Signal Process. 149 (2021) 107322.
[39] Jun Wu, Kui Hu, Yiwei Cheng, Haiping Zhu, Xinyu Shao, Yuanhang Wang, Data-driven remaining useful life prediction via multiple sensor signals and deep long short-term memory neural network, ISA Trans. 97 (2020) 241–250.
[40] Tangbin Xia, Ya Song, Yu Zheng, Ershun Pan, Lifeng Xi, An ensemble framework based on convolutional bi-directional LSTM with multiple time windows for remaining useful life estimation, Comput. Ind. 115 (2020) 103182.
[41] Xianping Zhong, Lin Zhang, Heng Ban, Deep reinforcement learning for class imbalance fault diagnosis of equipment in nuclear power plants, Ann. Nucl. Energy 184 (2023) 109685.
[42] Dhiraj Neupane, Jongwon Seok, Deep learning-based bearing fault detection using 2-D illustration of time sequence, in: 2020 International Conference on Information and Communication Technology Convergence, ICTC, 2020, pp. 562–566.
[43] Spyridon Plakias, Yiannis S. Boutalis, A novel information processing method based on an ensemble of auto-encoders for unsupervised fault detection, Comput. Ind. 142 (2022) 103743.
[44] Wade A. Smith, Robert B. Randall, Rolling element bearing diagnostics using the case western reserve university data: A benchmark study, Mech. Syst. Signal Process. 64 (2015) 100–131.
[45] Dhiraj Neupane, Yunsu Kim, Jongwon Seok, Bearing fault detection using scalogram and switchable normalization-based CNN (SN-CNN), IEEE Access 9 (2021) 88151–88166.

[46] Zhuyun Chen, Konstantinos Gryllias, Weihua Li, Intelligent fault diagnosis for rotary machinery using transferable convolutional neural network, IEEE Trans. Ind. Informat. 16 (1) (2020) 339–349.
[47] Changqing Shen, Jianye Zhou, Xinyu Yang, Lin Zhang, Siyu Shao, Gangying Bian, Multisignal VGG19 network with transposed convolution for rotating machinery fault diagnosis based on deep transfer learning, Shock. Vib. 2020 (2020) 8863388.
[48] Wentao Zhang, Ting Zhang, Guohua Cui, Ying Pan, Intelligent machine fault diagnosis using convolutional neural networks and transfer learning, IEEE Access 10 (2022) 50959–50973.
[49] Xudong Li, Yang Hu, Jianhua Zheng, Mingtao Li, Wenzhen Ma, Central moment discrepancy based domain adaptation for intelligent bearing fault diagnosis, Neurocomputing 429 (2021) 12–24.
[50] Jin Si, Hongmei Shi, Jingcheng Chen, Changchang Zheng, Unsupervised deep transfer learning with moment matching: A new intelligent fault diagnosis approach for bearings, Measurement 172 (2021) 108827.
[51] Cheng Cheng, Beitong Zhou, Guijun Ma, Dongrui Wu, Ye Yuan, Wasserstein distance based deep adversarial transfer learning for intelligent fault diagnosis with unlabeled or insufficient labeled data, Neurocomputing 409 (2020) 35–45.
[52] Qing Zhang, Lv Tang, Menglin Sun, Jianping Xuan, Tielin Shi, A statistical distribution recalibration method of soft labels to improve domain adaptation for cross-location and cross-machine fault diagnosis, Measurement 182 (2021) 109754.
[53] Na Lu, Tao Yin, Transferable common feature space mining for fault diagnosis with imbalanced data, Mech. Syst. Signal Process. 156 (2021) 107645.
[54] Zongliang Xie, Jinglong Chen, Yong Feng, Kaiyu Zhang, Zitong Zhou, End to end multi-task learning with attention for multi-objective fault diagnosis under small sample, J. Manuf. Syst. 62 (2022) 301–316.
[55] Xu Wang, Changqing Shen, Min Xia, Dong Wang, Jun Zhu, Zhongkui Zhu, Multi-scale deep intra-class transfer learning for bearing fault diagnosis, Reliab. Eng. Syst. Saf. 202 (2020) 107050.
[56] Tianci Zhang, Jinglong Chen, Fudong Li, Tongyang Pan, Shuilong He, A small sample focused intelligent fault diagnosis scheme of machines via multimodules learning with gradient penalized generative adversarial networks, IEEE Trans. Ind. Electron. 68 (10) (2020) 10130–10141.
[57] Changchang Che, Huawei Wang, Xiaomei Ni, Qiang Fu, Domain adaptive deep belief network for rolling bearing fault diagnosis, Comput. Ind. Eng. 143 (2020) 106427.
[58] Ansi Zhang, Shaobo Li, Yuxin Cui, Wanli Yang, Rongzhi Dong, Jianjun Hu, Limited data rolling bearing fault diagnosis with few-shot learning, Ieee Access 7 (2019) 110895–110904.
[59] Shaowei Liu, Hongkai Jiang, Zhenghong Wu, Xingqiu Li, Rolling bearing fault diagnosis using variational autoencoding generative adversarial networks with deep regret analysis, Measurement 168 (2021) 108371.
[60] Tianci Zhang, Jinglong Chen, Jingsong Xie, Tongyang Pan, SASLN: Signals augmented self-taught learning networks for mechanical fault diagnosis under small sample condition, IEEE Trans. Instrum. Meas. 70 (2021).
[61] Zhen Shi, Jinglong Chen, Yanyang Zi, Zitong Zhou, A novel multitask adversarial network via redundant lifting for multicomponent intelligent fault detection under sharp speed variation, IEEE Trans. Instrum. Meas. 70 (2021) 1–10.
[62] Haixin Lv, Jinglong Chen, Tongyang Pan, Zitong Zhou, Hybrid attribute conditional adversarial denoising autoencoder for zero-shot classification of mechanical intelligent fault diagnosis, Appl. Soft Comput. 95 (2020) 106577.
[63] Tongyang Pan, Jinglong Chen, Jingsong Xie, Zitong Zhou, Shuilong He, Deep feature generating network: A new method for intelligent fault detection of mechanical systems under class imbalance, IEEE Trans. Ind. Informat. 17 (9) (2020) 6282–6293.
[64] Tongyang Pan, Jinglong Chen, Cheng Qu, Zitong Zhou, A method for mechanical fault recognition with unseen classes via unsupervised convolutional adversarial auto-encoder, Meas. Sci. Technol. 32 (3) (2020) 035113.
[65] Tongyang Pan, Jinglong Chen, Jinsong Xie, Yuanhong Chang, Zitong Zhou, Intelligent fault identification for industrial automation system via multi-scale convolutional generative adversarial network with partially labeled samples, ISA Trans. 101 (2020) 379–389.
[66] Pengfei Liang, Chao Deng, Jun Wu, Guoqiang Li, Zhixin Yang, Yuanhang Wang, Intelligent fault diagnosis via semisupervised generative adversarial nets and wavelet transform, IEEE Trans. Instrum. Meas. 69 (7) (2019) 4659–4671.
[67] Xiang Li, Wei Zhang, Hui Ma, Zhong Luo, Xu Li, Partial transfer learning in machinery cross-domain fault diagnostics using class-weighted adversarial networks, Neural Netw. 129 (2020) 313–322.
[68] Behnoush Rezaeianjouybari, Yi Shang, A novel deep multi-source domain adaptation framework for bearing fault diagnosis based on feature-level and task-specific distribution alignment, Measurement 178 (2021) 109359.
[69] Jun Zhu, Nan Chen, Changqing Shen, A new multiple source domain adaptation fault diagnosis method between different rotating machines, IEEE Trans. Ind. Informat. 17 (7) (2020) 4788–4797.
[70] Yu Ding, Liang Ma, Jian Ma, Mingliang Suo, Laifa Tao, Yujie Cheng, Chen Lu, Intelligent fault diagnosis for rotating machinery using deep Q-network based health state classification: A deep reinforcement learning approach, Adv. Eng. Informat. 42 (2019) 100977.
[71] Gensheng Qian, Jingquan Liu, Development of deep reinforcement learning-based fault diagnosis method for rotating machinery in nuclear power plants, Prog. Nucl. Energy 152 (2022) 104401.
[72] Jian Zhou, Lianyu Zheng, Yiwei Wang, Cheng Wang, Robert X. Gao, Automated model generation for machinery fault diagnosis based on reinforcement learning and neural architecture search, IEEE Trans. Instrum. Meas. 71 (2022) 1–12.
[73] Siyu Shao, Stephen McAleer, Ruqiang Yan, Pierre Baldi, Highly accurate machine fault diagnosis using deep transfer learning, IEEE Trans. Ind. Informat. 15 (4) (2018) 2446–2455.
[74] Ruixin Wang, Hongkai Jiang, Xingqiu Li, Shaowei Liu, A reinforcement neural architecture search method for rolling bearing fault diagnosis, Measurement 154 (2020) 107417.
[75] Zisheng Wang, Jianping Xuan, Tielin Shi, Multi-label fault recognition framework using deep reinforcement learning and curriculum learning mechanism, Adv. Eng. Informat. 54 (2022) 101773.
[76] Wei Zhang, Chuanhao Li, Gaoliang Peng, Yuanhang Chen, Zhujun Zhang, A deep convolutional neural network with new training methods for bearing fault diagnosis under noisy environment and different working load, Mech. Syst. Signal Process. 100 (2018) 439–453.
[77] Siyu Zhang, Hua-Liang Wei, Jinliang Ding, An effective zero-shot learning approach for intelligent fault detection using 1D CNN, Appl. Intell. 53 (12) (2023) 16041–16058.
[78] Qingzhe Wei, Xincheng Tian, Long Cui, Fuquan Zheng, Lida Liu, WSAFormer-DFFN: A model for rotating machinery fault diagnosis using 1d window-based multi-head self-attention and deep feature fusion network, Eng. Appl. Artif. Intell. 124 (2023) 106633.
[79] Yuhong Jin, Lei Hou, Yushu Chen, A time series transformer based method for the rotating machinery fault diagnosis, Neurocomputing 494 (2022) 379–395.
[80] Xiao Yu, Zhongting Liang, Youjie Wang, Hongshen Yin, Xiaowen Liu, Wanli Yu, Yanqiu Huang, A wavelet packet transform-based deep feature transfer learning method for bearing fault diagnosis under different working conditions, Measurement 201 (2022) 111597.
[81] Zhiyong Luo, Yueyue Peng, Xin Dong, Hao Qian, Rotating machinery fault diagnosis using dimension expansion and AntisymNet lightweight convolutional neural network, Meas. Sci. Technol. 34 (11) (2023) 115005.
[82] Jurgen Van Den Hoogen, Dan Hudson, Stefan Bloemheuvel, Martin Atzmueller, Hyperparameter analysis of wide-kernel CNN architectures in industrial fault detection: an exploratory study, Int. J. Data Sci. Anal. (2023) 1–22.
[83] Yuling Tian, Xiangyu Liu, A deep adaptive learning method for rolling bearing fault diagnosis using immunity, Tsinghua Sci. Technol. 24 (6) (2019) 750–762.
[84] Tian Han, Ruiyi Ma, Jigui Zheng, Combination bidirectional long short-term memory and capsule network for rotating machinery fault diagnosis, Measurement 176 (2021) 109208.
[85] DeCai Li, Mian Zhang, TianBo Kang, Bin Li, HongBiao Xiang, KeSheng Wang, ZhenLong Pei, XiuYang Tang, Peng Wang, Fault diagnosis of rotating machinery based on dual convolutional-capsule network (DC-CN), Measurement 187 (2022) 110258.
[86] Ruyi Huang, Jipu Li, Shuhua Wang, Guanghui Li, Weihua Li, A robust weight-shared capsule network for intelligent machinery fault diagnosis, IEEE Trans. Ind. Informat. 16 (10) (2020) 6466–6475.
[87] Yuantao Yang, Huailiang Zheng, Yongbo Li, Minqiang Xu, Yushu Chen, A fault diagnosis scheme for rotating machinery using hierarchical symbolic analysis and convolutional neural network, ISA Trans. 91 (2019) 235–252.
[88] Yiping Gao, Liang Gao, Xinyu Li, Siyu Cao, A hierarchical training-convolutional neural network for imbalanced fault diagnosis in complex equipment, IEEE Trans. Ind. Informat. 18 (11) (2022) 8138–8145.
[89] Zifei Xu, Xuan Mei, Xinyu Wang, Minnan Yue, Jiangtao Jin, Yang Yang, Chun Li, Fault diagnosis of wind turbine bearing using a multi-scale convolutional neural network with bidirectional long short term memory and weighted majority voting for multi-sensors, Renew. Energy 182 (2022) 615–626.
[90] Shihang Yu, Min Wang, Shanchen Pang, Limei Song, Xue Zhai, Yawu Zhao, TDMSAE: A transferable decoupling multi-scale autoencoder for mechanical fault diagnosis, Mech. Syst. Signal Process. 185 (2023) 109789.
[91] Syahril Ramadhan Saufi, Muhammad Firdaus Isham, Zair Asrar Ahmad, M. Danial Abu Hasan, Machinery fault diagnosis based on a modified hybrid deep sparse autoencoder using a raw vibration time-series signal, J. Ambient. Intell. Humaniz. Comput. 14 (4) (2023) 3827–3838.
[92] Jiahui Tang, Jimei Wu, Bingbing Hu, Jie Liu, Towards a fault diagnosis method for rolling bearing with bi-directional deep belief network, Appl. Acoust. 192 (2022) 108727.
[93] Huimin Zhao, Jie Liu, Huayue Chen, Jie Chen, Yang Li, Junjie Xu, Wu Deng, Intelligent diagnosis using continuous wavelet transform and gauss convolutional deep belief network, IEEE Trans. Reliab. (2022).
[94] Shuzhi Gao, Lintao Xu, Yimin Zhang, Zhiming Pei, Rolling bearing fault diagnosis based on SSA optimized self-adaptive DBN, ISA Trans. 128 (2022) 485–502.
[95] Tianchen Liang, Shuaipeng Wu, Wenjing Duan, Ruishan Zhang, Bearing fault diagnosis based on improved ensemble learning and deep belief network, in: J. Phys. Conf. Series, 1074, IOP Publishing, 2018, 012154.

[96] Xiaoan Yan, Ying Liu, Minping Jia, Multiscale cascading deep belief network for fault identification of rotating machinery under various working conditions, Knowl.-Based Syst. 193 (2020) 105484.
[97] Xingqing Wang, Yanfeng Li, Ting Rui, Huijie Zhu, Jianchao Fei, Bearing fault diagnosis method based on Hilbert envelope spectrum and deep belief network, J. Vibroeng. 17 (3) (2015) 1295–1308.
[98] Yifei Ding, Minping Jia, Jichao Zhuang, Yudong Cao, Xiaoli Zhao, Chi-Guhn Lee, Deep imbalanced domain adaptation for transfer learning fault diagnosis of bearings under multiple working conditions, Reliab. Eng. Syst. Saf. 230 (2023) 108890.
[99] Liang Guo, Yaguo Lei, Saibo Xing, Tao Yan, Naipeng Li, Deep convolutional transfer learning network: A new method for intelligent fault diagnosis of machines with unlabeled data, IEEE Trans. Ind. Electron. 66 (9) (2019) 7316–7325.
[100] Alessandro Paolo Daga, Alessandro Fasana, Stefano Marchesiello, Luigi Garibaldi, The politecnico di torino rolling bearing test rig: Description and analysis of open access data, Mech. Syst. Signal Process. 120 (2019) 252–273.
[101] Huan Wang, Zhiliang Liu, Dandan Peng, Ming J. Zuo, Interpretable convolutional neural network with multilayer wavelet for noise-robust machinery fault diagnosis, Mech. Syst. Signal Process. 195 (2023) 110314.
[102] Amir Hosein Zamanian, Experimental dataset for gear fault diagnosis, 2014.
[103] Patrick Nectoux, Rafael Gouriveau, Kamal Medjaher, Emmanuel Ramasso, Brigitte Chebel-Morello, Noureddine Zerhouni, Christophe Varnier, PRONOSTIA : An experimental platform for bearings accelerated degradation tests., in: IEEE International Conference on Prognostics and Health Management, PHM'12., Vol. sur CD ROM, IEEE Catalog Number : CPF12PHM-CDR, Denver, Colorado, United States, 2012, pp. 1–8.
[104] Lu Liu, Xiao Song, Kai Chen, Baocun Hou, Xudong Chai, Huansheng Ning, An enhanced encoder–decoder framework for bearing remaining useful life prediction, Measurement 170 (2021) 108753.
[105] Lei Nie, Lvfan Zhang, Shiyi Xu, Wentao Cai, Haoming Yang, Remaining useful life prediction for rolling bearings based on similarity feature fusion and convolutional neural network, J. Braz. Soc. Mech. Sci. Eng. 44 (8) (2022) 328.
[106] Chang-Hua Hu, Hong Pei, Xiao-Sheng Si, Dang-Bo Du, Zhe-Nan Pang, Xi Wang, A prognostic model based on DBN and diffusion process for degrading bearing, IEEE Trans. Ind. Electron. 67 (10) (2020) 8767–8777.
[107] Cédric Peeters, Wenyi Wang, David Blunt, Timothy Verstraeten, Jan Helsen, Fatigue crack detection in planetary gears: Insights from the HUMS2023 data challenge, Mech. Syst. Signal Process. 212 (2024) 111292.
[108] Nader Sawalhi, Wenyi Wang, David Blunt, Helicopter planet gear rim crack diagnosis and trending using cepstrum editing enhanced with deconvolution, 2024.
[109] Jun Dai, Jun Wang, Weiguo Huang, Changqing Shen, Juanjuan Shi, Xingxing Jiang, Zhongkui Zhu, Bearing anomaly detection based on generative adversarial network, in: Advances in Condition Monitoring and Structural Health Monitoring: WCCM 2019, Springer, 2021, pp. 207–215.
[110] Yongchao Zhang, Zhaohui Ren, Shihua Zhou, An intelligent fault diagnosis for rolling bearing based on adversarial semi-supervised method, IEEE Access 8 (2020) 149868–149877.
[111] Levent Eren, Bearing fault detection by one-dimensional convolutional neural networks, Math. Probl. Eng. 2017 (2017) 1–9.
[112] Ibrahim Halil Ozcan, Ozer Can Devecioglu, Turker Ince, Levent Eren, Murat Askar, Enhanced bearing fault detection using multichannel, multilevel 1D CNN classifier, Electr. Eng. 104 (2) (2022) 435–447.
[113] Mohamad Najafi, Yasser Baleghi, Sayyed Asghar Gholamian, Seyyed Mehdi Mirimani, Fault diagnosis of electrical equipment through thermal imaging and interpretable machine learning applied on a newly-introduced dataset, in: 2020 6th Iranian Conference on Signal Processing and Intelligent Systems, ICSPIS, 2020, pp. 1–7.
[114] Lifu Xu, Soo Siang Teoh, Haidi Ibrahim, A deep learning approach for electric motor fault diagnosis based on modified InceptionV3, Sci. Rep. 14 (1) (2024) 12344.
[115] Qin Zhao, Yu Ding, Chen Lu, Chao Wang, Liang Ma, Laifa Tao, Jian Ma, An adaptive fault diagnosis framework under class-imbalanced conditions based on contrastive augmented deep reinforcement learning, Expert Syst. Appl. 234 (2023) 121001.
[116] Zhibin Zhao, Tianfu Li, Jingyao Wu, Chuang Sun, Shibin Wang, Ruqiang Yan, Xuefeng Chen, Deep learning algorithms for rotating machinery intelligent diagnosis: An open source benchmark study, ISA Trans. 107 (2020) 224–255.
[117] Guoqiang Li, Jun Wu, Chao Deng, Xuebing Xu, Xinyu Shao, Deep reinforcement learning-based online domain adaptation method for fault diagnosis of rotating machinery, IEEE/ASME Trans. Mechatronics 27 (5) (2022) 2796–2805.
[118] Yiping Gao, Liang Gao, Xinyu Li, Xuguo Yan, A semi-supervised convolutional neural network-based method for steel surface defect recognition, Robot. Comput.-Integr. Manuf. 61 (2020) 101825.
[119] Yucheng Wang, Liang Gao, Yiping Gao, Xinyu Li, A new graph-based semi-supervised method for surface defect classification, Robot. Comput.-Integr. Manuf. 68 (2021) 102083.
[120] Shawn. Sheng, Wind turbine gearbox condition monitoring vibration analysis benchmarking datasets, 2014.
[121] Pengfei Liang, Chao Deng, Jun Wu, Zhixin Yang, Jinxuan Zhu, Zihan Zhang, Single and simultaneous fault diagnosis of gearbox via a semi-supervised and high-accuracy adversarial learning framework, Knowl.-Based Syst. 198 (2020) 105895.
[122] Yunhan Kim, Kyumin Na, Byeng D. Youn, A health-adaptive time-scale representation (HTSR) embedded convolutional neural network for gearbox fault diagnostics, Mech. Syst. Signal Process. 167 (2022) 108575.
[123] Luyang Jing, Ming Zhao, Pin Li, Xiaoqiang Xu, A convolutional neural network based feature learning and fault diagnosis method for the condition monitoring of gearbox, Measurement 111 (2017) 1–10.
[124] Tam T. Truong, Jay Airao, Faramarz Hojati, Charlotte F. Ilvig, Bahman Azarhoushang, Panagiotis Karras, Ramin Aghababaei, Data-driven prediction of tool wear using bayesian regularized artificial neural networks, Measurement 238 (2024) 115303.
[125] Botao Wang, Chuanwen Shen, Chenxi Yu, Yutao Yang, Data fused motor fault identification based on adversarial auto-encoder, in: 2019 IEEE 10th International Symposium on Power Electronics for Distributed Generation Systems, PEDG, IEEE, 2019, pp. 299–305.
[126] Jing Yang, Guo Xie, Yanxi Yang, Fault diagnosis of rotating machinery using denoising-integrated sparse autoencoder based health state classification, IEEE Access 11 (2023) 15174–15183.
[127] Jiahui Tang, Jimei Wu, Jiajuan Qing, A feature learning method for rotating machinery fault diagnosis via mixed pooling deep belief network and wavelet transform, Results Phys. 39 (2022) 105781.
[128] Dong Liu, Zhihuai Xiao, Xiao Hu, Congxin Zhang, OP Malik, Feature extraction of rotor fault based on EEMD and curve code, Measurement 135 (2019) 712–724.
[129] Xiao Hu, Daohong Wei, Dong Liu, Zhihuai Xiao, Xiangchen Xia, O.P. Malik, Fault diagnosis for rotor based on multi-sensor information and progressive strategies, Meas. Sci. Technol. 34 (6) (2023) 065111.
[130] Jingyao Wu, Zhibin Zhao, Chuang Sun, Ruqiang Yan, Xuefeng Chen, Fault-attention generative probabilistic adversarial autoencoder for machine anomaly detection, IEEE Trans. Ind. Informat. 16 (12) (2020) 7479–7488.
[131] Hui Wang, Jiawen Xu, Chuang Sun, Ruqiang Yan, Xuefeng Chen, Intelligent fault diagnosis for planetary gearbox using time-frequency representation and deep reinforcement learning, IEEE/ASME Trans. Mechatronics 27 (2) (2022) 985–998.
[132] Ting Huang, Qiang Zhang, Xiaoan Tang, Shuangyao Zhao, Xiaonong Lu, A novel fault diagnosis method based on CNN and LSTM and its application in fault diagnosis for complex systems, Artif. Intell. Rev. (2022) 1–27.
[133] Peng Peng, Wenjia Zhang, Yi Zhang, Yanyan Xu, Hongwei Wang, Heming Zhang, Cost sensitive active learning using bidirectional gated recurrent neural networks for imbalanced fault diagnosis, Neurocomputing 407 (2020) 232–245.
[134] Yalin Wang, Zhuofu Pan, Xiaofeng Yuan, Chunhua Yang, Weihua Gui, A novel deep learning based fault diagnosis approach for chemical process with extended deep belief network, ISA Trans. 96 (2020) 457–467.
[135] Meng Rao, Ming J. Zuo, Zhigang Tian, A speed normalized autoencoder for rotating machinery fault detection under varying speed conditions, Mech. Syst. Signal Process. 189 (2023) 110109.
[136] Dongdong Wei, Te Han, Fulei Chu, Ming Jian Zuo, Weighted domain adaptation networks for machinery fault diagnosis, Mech. Syst. Signal Process. 158 (2021) 107744.
[137] Meng Rao, Qing Li, Dongdong Wei, Ming J. Zuo, A deep bi-directional long short-term memory model for automatic rotating speed extraction from raw vibration signals, Measurement 158 (2020) 107719.
[138] Yuejian Chen, Xihui Liang, Ming J. Zuo, Sparse time series modeling of the baseline vibration from a gearbox under time-varying speed condition, Mech. Syst. Signal Process. 134 (2019) 106342.
[139] Yuejian Chen, Xihui Liang, Ming J. Zuo, An improved singular value decomposition-based method for gear tooth crack detection and severity assessment, J. Sound Vib. 468 (2020) 115068.
[140] Zhongliang Chen, Zhuofei Huang, Wenxiong Kang, A novel transfer learning method utilizing acoustic and vibration signals for rotating machinery fault diagnosis, 2023, arXiv preprint arXiv:2310.14796.
[141] Yuanyuan Li, Lingyu Hou, Ming Tang, Qichun Sun, Jiahang Chen, Wanqing Song, Wei Yao, Le Cao, Prediction of wind turbine blades icing based on feature selection and 1D-CNN-SBiGRU, Multimedia Tools Appl. 81 (3) (2022) 4365–4385.
[142] Jiaohui Xu, Wen Tan, Tingshun Li, Predicting fan blade icing by using particle swarm optimization and support vector machine algorithm, Comput. Electr. Eng. 87 (2020) 106751.
[143] Jin Xiao, Chunyan Li, Bo Liu, Jing Huang, Ling Xie, Prediction of wind turbine blade icing fault based on selective deep ensemble model, Knowl.-Based Syst. 242 (2022) 108290.
[144] Huajin Li, Jiahao Deng, Shuang Yuan, Peng Feng, Dimuthu D.K. Arachchige, Monitoring and identifying wind turbine generator bearing faults using deep belief network and EWMA control charts, Front. Energy Res. 9 (2021) 799039.
[145] Welson Bassi, Alcantaro Lemes Rodrigues, Ildo Luis Sauer, Dataset on SCADA data of an urban small wind turbine operation in sao paulo, Brazil, Data 8 (3) (2023).

[146] Upma Singh, M. Rizwan, SCADA system dataset exploration and machine learning based forecast for wind turbines, Results Eng. 16 (2022) 100640.
[147] Yong Duan, Xiangang Cao, Jiangbin Zhao, Man Li, Xin Yang, A spatiotemporal fusion autoencoder-based health indicator automatic construction method for rotating machinery considering vibration signal expression, IEEE Sensors J. 23 (20) (2023) 24822–24838.
[148] Xiaoan Yan, Daoming She, Yadong Xu, Deep order-wavelet convolutional variational autoencoder for fault identification of rolling bearing under fluctuating speed conditions, Expert Syst. Appl. 216 (2023) 119479.
[149] Wenxin Dai, Zhenling Mo, Chong Luo, Jing Jiang, Heng Zhang, Qiang Miao, Fault diagnosis of rotating machinery based on deep reinforcement learning and reciprocal of smoothness index, IEEE Sensors J. 20 (15) (2020) 8307–8315.
[150] Sheng Guo, Bin Zhang, Tao Yang, Dongzhen Lyu, Wei Gao, Multitask convolutional neural network with information fusion for bearing fault diagnosis and localization, IEEE Trans. Ind. Electron. 67 (9) (2019) 8005–8015.
[151] Z. Peng, N.J. Kessissoglou, M. Cox, A study of the effect of contaminant particles in lubricants using wear debris and vibration condition monitoring techniques, Wear 258 (11) (2005) 1651–1662.
[152] Peng Chen, Yu Li, Kesheng Wang, Ming J. Zuo, A novel knowledge transfer network with fluctuating operational condition adaptation for bearing fault pattern recognition, Measurement 158 (2020) 107739.
[153] Shuai Yang, Yifei Zhou, Xu Chen, Chuan Li, Heng Song, Fault diagnosis for wind turbines with graph neural network model based on one-shot learning, R. Soc. Open Sci. 10 (7) (2023) 230706.
[154] Juan Xu, Long Zhou, Weihua Zhao, Yuqi Fan, Xu Ding, Xiaohui Yuan, Zero-shot learning for compound fault diagnosis of bearings, Expert Syst. Appl. 190 (2022) 116197.
[155] Chuan Li, Diego Cabrera, Fernando Sancho, René-Vinicio Sánchez, Mariela Cerrada, Jianyu Long, José Valente de Oliveira, Fusing convolutional generative adversarial encoders for 3D printer fault detection with only normal condition signals, Mech. Syst. Signal Process. 147 (2021) 107108.
[156] Yixiao Liao, Ruyi Huang, Jipu Li, Zhuyun Chen, Weihua Li, Deep semisupervised domain generalization network for rotary machinery fault diagnosis under variable speed, IEEE Trans. Instrum. Meas. 69 (10) (2020) 8064–8075.
[157] Jinyang Jiao, Jing Lin, Ming Zhao, Kaixuan Liang, Double-level adversarial domain adaptation network for intelligent fault diagnosis, Knowl.-Based Syst. 205 (2020) 106236.
[158] Donatella Zappalá, Nur Sarma, S. Djurović, C.J. Crabtree, A. Mohammad, P.J. Tavner, Electrical & mechanical diagnostic indicators of wind turbine induction generator rotor faults, Renew. Energy 131 (2019) 14–24.
[159] Zisheng Wang, Qing Zhang, Lv Tang, Tielin Shi, Jianping Xuan, Transfer reinforcement learning method with multi-label learning for compound fault recognition, Adv. Eng. Informat. 55 (2023) 101818.
[160] Dajian Huang, Wen-An Zhang, Fanghong Guo, Weijiang Liu, Xiaoming Shi, Wavelet packet decomposition-based multiscale CNN for fault diagnosis of wind turbine gearbox, IEEE Trans. Cybern. 53 (1) (2023) 443–453.
[161] Ruo-Bin Sun, Zhi-Bo Yang, Zhi Zhai, Xue-Feng Chen, Sparse representation based on parametric impulsive dictionary design for bearing fault diagnosis, Mech. Syst. Signal Process. 122 (2019) 737–753.
[162] Sujesh Kumar, M. Lokesha, Kiran Kumar, K.R. Srinivas, Vibration based fault diagnosis techniques for rotating mechanical components: Review paper, IOP Conf. Series: Mater. Sci. Eng. 376 (1) (2018) 012109.
[163] R. Collacott, Mechanical fault diagnosis and condition monitoring, Springer Science & Business Media, 2012.
[164] Salah M. Ali Al-Obaidi, M. Salman Leong, R.I. Raja Hamzah, Ahmed M. Abdelrhman, A review of acoustic emission technique for machinery condition monitoring: defects detection & diagnostic, Appl. Mech. Mater. 229 (2012) 1476–1480.
[165] Tauheed Mian, Anurag Choudhary, Shahab Fatima, Vibration and infrared thermography based multiple fault diagnosis of bearing using deep learning, Nondestruct. Test. Eval. 38 (2) (2023) 275–296.
[166] Hanyang Liu, Feng Sheng, Xinyu Lu, Jing Lin, MCSA-based fault diagnosis for rotary parts in servo motion system: Approach and experimental study, in: J. Phys. Conf. Series, 2762, (1) IOP Publishing, 2024, 012055.
[167] Olivier Janssens, Viktor Slavkovikj, Bram Vervisch, Kurt Stockman, Mia Loccufier, Steven Verstockt, Rik Van de Walle, Sofie Van Hoecke, Convolutional neural network based fault detection for rotating machinery, J. Sound Vib. 377 (2016) 331–345.
[168] Rui Zhao, Ruqiang Yan, Jinjiang Wang, Kezhi Mao, Learning to monitor machine health with convolutional bi-directional LSTM networks, Sensors 17 (2) (2017).
[169] Ignacio Martin-Diaz, Daniel Morinigo-Sotelo, Oscar Duque-Perez, Rene de J. Romero-Troncoso, Early fault detection in induction motors using AdaBoost with imbalanced small data and optimized sampling, IEEE Trans. Ind. Appl. 53 (3) (2017) 3066–3075.
[170] Jianan Wei, Haisong Huang, Liguo Yao, Yao Hu, Qingsong Fan, Dong Huang, New imbalanced bearing fault diagnosis method based on sample-characteristic oversampling technique (SCOTE) and multi-class LS-SVM, Appl. Soft Comput. 101 (2021) 107043.
[171] Wei Hao, Feng Liu, Imbalanced data fault diagnosis based on an evolutionary online sequential extreme learning machine, Symmetry 12 (8) (2020) 1204.
[172] Chuanxia Jian, Yinhui Ao, Imbalanced fault diagnosis based on semi-supervised ensemble learning, J. Intell. Manuf. (2022) 1–16.
[173] Roozbeh Razavi-Far, Maryam Farajzadeh-Zanjani, Mehrdad Saif, An integrated class-imbalanced learning scheme for diagnosing bearing defects in induction motors, IEEE Trans. Ind. Informat. 13 (6) (2017) 2758–2769.
[174] Jianan Wei, Haisong Huang, Liguo Yao, Yao Hu, Qingsong Fan, Dong Huang, New imbalanced fault diagnosis framework based on cluster-MWMOTE and MFO-optimized LS-SVM using limited and complex bearing data, Eng. Appl. Artif. Intell. 96 (2020) 103966.
[175] Wei Zhang, Gaoliang Peng, Chuanhao Li, Yuanhang Chen, Zhujun Zhang, A new deep learning model for fault diagnosis with good anti-noise and domain adaptation ability on raw vibration signals, Sensors 17 (2) (2017).
[176] Ruxue Bai, Quansheng Xu, Zong Meng, Lixiao Cao, Kangshuo Xing, Fengjie Fan, Rolling bearing fault diagnosis based on multi-channel convolution neural network and multi-scale clipping fusion data augmentation, Measurement 184 (2021) 109885.
[177] Xiang Li, Wei Zhang, Qian Ding, Jian-Qiao Sun, Intelligent rotating machinery fault diagnosis based on deep learning using data augmentation, J. Intell. Manuf. 31 (2) (2020) 433–452.
[178] Haixin Lv, Jinglong Chen, Tianci Zhang, Rujie Hou, Tongyang Pan, Zitong Zhou, SDA: Regularization with cut-flip and mix-normal for machinery fault diagnosis under small dataset, ISA Trans. 111 (2021) 337–349.
[179] Jin Woo Oh, Jongpil Jeong, Data augmentation for bearing fault detection with a light weight CNN, Procedia Comput. Sci. 175 (2020) 72–79.
[180] Jose-Luis Casteleiro-Roca, Seungmin Han, Seokju Oh, Jongpil Jeong, Bearing fault diagnosis based on multiscale convolutional neural network using data augmentation, J. Sensors 2021 (2021) 6699637.
[181] Jinrui Wang, Baokun Han, Huaiqian Bao, Mingyan Wang, Zhenyun Chu, Yuwei Shen, Data augment method for machine fault diagnosis using conditional generative adversarial networks, Proc. Inst. Mech. Eng. Part D: J. Automob. Eng. 234 (12) (2020) 2719–2727.
[182] Rugen Wang, Zhuyun Chen, Weihua Li, Gradient flow-based meta generative adversarial network for data augmentation in fault diagnosis, Appl. Soft Comput. 142 (2023) 110313.
[183] Yunpeng Liu, Hongkai Jiang, Chaoqiang Liu, Wangfeng Yang, Wei Sun, Data-augmented wavelet capsule generative adversarial network for rolling bearing fault diagnosis, Knowl.-Based Syst. 252 (2022) 109439.
[184] Penghui Zhao, Qinghe Zheng, Zhongjun Ding, Yi Zhang, Hongjun Wang, Yang Yang, A high-dimensional and small-sample submersible fault detection method based on feature selection and data augmentation, Sensors 22 (1) (2022).
[185] Xin Gao, Fang Deng, Xianghu Yue, Data augmentation in fault diagnosis based on the wasserstein generative adversarial network with gradient penalty, Neurocomputing 396 (2020) 487–494.
[186] Changchang Che, Huawei Wang, Ruiguan Lin, Xiaomei Ni, Deep meta-learning and variational autoencoder for coupling fault diagnosis of rolling bearing under variable working conditions, Proc. Inst. Mech. Eng. Part C: J. Mech. Eng. Sci. 236 (17) (2022) 9900–9913.
[187] You ren Wang, Guo dong Sun, Qi Jin, Imbalanced sample fault diagnosis of rotating machinery using conditional variational auto-encoder generative adversarial network, Appl. Soft Comput. 92 (2020) 106333.
[188] He Li, Zhijin Zhang, Chunlei Zhang, Data augmentation via variational mode reconstruction and its application in few-shot fault diagnosis of rolling bearings, Measurement 217 (2023) 113062.
[189] Jilun Tian, Yuchen Jiang, Jiusi Zhang, Hao Luo, Shen Yin, A novel data augmentation approach to fault diagnosis with class-imbalance problem, Reliab. Eng. Syst. Saf. 243 (2024) 109832.
[190] Rolf Isermann, Model-based fault-detection and diagnosis–status and applications, Annu. Rev. Control. 29 (1) (2005) 71–85.
[191] Zhipeng Feng, Ming Liang, Fulei Chu, Recent advances in time–frequency analysis methods for machinery fault diagnosis: A review with application examples, Mech. Syst. Signal Process. 38 (1) (2013) 165–205.
[192] M. Amine Atoui, Achraf Cohen, Coupling data-driven and model-based methods to improve fault diagnosis, Comput. Ind. 128 (2021) 103401.
[193] Jae Hong Suh, Soundar R.T. Kumara, Shreesh P. Mysore, Machinery fault diagnosis and prognosis: application of advanced signal processing techniques, CIRP Ann 48 (1) (1999) 317–320.
[194] M. Siva Santhoshi, K. Sharath Babu, Sanjeev Kumar, Durgesh Nandan, An investigation on rolling element bearing fault and real-time spectrum analysis by using short-time fourier transform, in: Proceedings of International Conference on Recent Trends in Machine Learning, IoT, Smart Cities and Applications: ICMISC 2020, Springer, 2021, pp. 561–567.
[195] Mehrdad Heydarzadeh, Mohsen Zafarani, Mehrdad Nourani, Bilal Akin, A wavelet-based fault diagnosis approach for permanent magnet synchronous motors, IEEE Trans. Energy Convers. 34 (2) (2018) 761–772.
[196] Jinde Zheng, Junsheng Cheng, Yu Yang, Generalized empirical mode decomposition and its applications to rolling element bearing fault diagnosis, Mech. Syst. Signal Process. 40 (1) (2013) 136–153.

[197] Zoltan German-Sallo, Horatiu Stefan Grif, Hilbert-huang transform in fault detection, Procedia Manuf. 32 (2019) 591–595, 12th International Conference Interdisciplinarity in Engineering, INTER-ENG 2018, 4–5 October 2018, Tirgu Mures, Romania.

[198] Hans Konstantin-Hansen, Henrik Herlufsen, Envelope and cepstrum analyses for machinery fault identification, Sound & Vib. 44 (5) (2010) 10–12.

[199] Min-Chun Pan, Wen-Chang Tsao, Using appropriate IMFs for envelope analysis in multiple fault diagnosis of ball bearings, Int. J. Mech. Sci. 69 (2013) 114–124.

[200] Xian-Bo Wang, Zhi-Xin Yang, Xiao-An Yan, Novel particle swarm optimization-based variational mode decomposition method for the fault diagnosis of complex rotating machinery, IEEE/ASME Trans. Mechatronics 23 (1) (2018) 68–79.

[201] A. Malhi, R.X. Gao, PCA-based feature selection scheme for machine defect classification, IEEE Trans. Instrum. Meas. 53 (6) (2004) 1517–1525.

[202] Zhipeng Ma, Ming Zhao, Shuai Chen, Dong Guo, Encoder-based weak fault detection for rotating machinery using improved Gaussian process regression, Struct. Heal. Monit. 20 (1) (2021) 255–272.

[203] Zhigang Tian, Lorna Wong, Nima Safaei, A neural network approach for remaining useful life prediction utilizing both failure and suspension histories, Mech. Syst. Signal Process. 24 (5) (2010) 1542–1555.

[204] Lifeng Wu, Beibei Yao, Zhen Peng, Yong Guan, Fault diagnosis of roller bearings based on a wavelet neural network and manifold learning, Appl. Sci. 7 (2) (2017) 158.

[205] Hosameldin Ahmed, Asoke K. Nandi, Compressive sampling and feature ranking framework for bearing fault classification with vibration signals, IEEE Access 6 (2018) 44731–44746.

[206] Tao Tang, Lin Bo, Xiaofeng Liu, Bing Sun, Daiping Wei, Variable predictive model class discrimination using novel predictive models and adaptive feature selection for bearing fault identification, J. Sound Vib. 425 (2018) 137–148.

[207] Fan Jiang, Zhencai Zhu, Wei Li, Gongbo Zhou, Shixiong Xia, Noise reduction in feature level and its application in rolling element bearing fault diagnosis, Adv. Mech. Eng. 10 (3) (2018) 1687814018764820.

[208] Sunil Tyagi, S.K. Panigrahi, A hybrid genetic algorithm and back-propagation classifier for gearbox fault diagnosis, Appl. Artif. Intell. 31 (7–8) (2017) 593–612.

[209] Tayyab Waqar, Mustafa Demetgul, Thermal analysis MLP neural network based fault diagnosis on worm gears, Measurement 86 (2016) 56–66.

[210] Hanxin Chen, Yanjun Lu, Ling Tu, Fault identification of gearbox degradation with optimized wavelet neural network, Shock. Vib. 20 (2) (2013) 247–262.

[211] S. Saeid Moosavi, Abdesslem Djerdir, Y. Ait-Amirat, Davod Arab Khaburi, ANN based fault diagnosis of permanent magnet synchronous motor under stator winding shorted turn, Electr. Power Syst. Res. 125 (2015) 67–82.

[212] M. Hernandez-Vargas, Eduardo Cabal-Yepez, Arturo Garcia-Perez, Real-time SVD-based detection of multiple combined faults in induction motors, Comput. Electr. Eng. 40 (7) (2014) 2193–2203.

[213] Tahar Boukra, Abdesselam Lebaroud, Guy Clerc, Statistical and neural-network approaches for the classification of induction machine faults using the ambiguity plane representation, IEEE Trans. Ind. Electron. 60 (9) (2012) 4034–4042.

[214] Srinivas Chikkam, Sachin Singh, Condition monitoring and fault diagnosis of induction motor using dwt and ANN, Arab. J. Sci. Eng. 48 (5) (2023) 6237–6252.

[215] Mushabi Pule, Oduetse Matsebe, Ravi Samikannu, Application of PCA and SVM in fault detection and diagnosis of bearings with varying speed, Math. Probl. Eng. 2022 (1) (2022) 5266054.

[216] Abu Hanif Md Ripon, Muhammad Ahsan Ullah, Rotating machine fault detection using support vector machine (svm) classifier, in: 2023 4th International Conference on Computing and Communication Systems (I3CS), IEEE, 2023, pp. 1–5.

[217] Anupam Kumar, Anand Parey, Pavan Kumar Kankar, Supervised machine learning based approach for early fault detection in polymer gears using vibration signals, MAPAN 38 (2) (2023) 383–394.

[218] M. Kordestani, M. Rezamand, M. Orchard, R. Carriveau, D.S.K. Ting, M. Saif, Planetary gear faults detection in wind turbine gearbox based on a ten years historical data from three wind farms, IFAC- Pap. 53 (2) (2020) 10318–10323.

[219] I.M. Jamadar, R. Nithin, S. Nagashree, V.R. Prajwal Prasad, M. Preetham, P.K. Samal, Shekhar Singh, Spur gear fault detection using design of experiments and support vector machine (SVM) algorithm, J. Fail. Anal. Prev. 23 (5) (2023) 2014–2028.

[220] Orhan Yaman, An automated faults classification method based on binary pattern and neighborhood component analysis using induction motor, Measurement 168 (2021) 108323.

[221] Min-Chan Kim, Jong-Hyun Lee, Dong-Hun Wang, In-Soo Lee, Induction motor fault diagnosis using support vector machine, neural networks, and boosting methods, Sensors 23 (5) (2023) 2585.

[222] Jing Tian, Carlos Morillo, Michael H Azarian, Michael Pecht, Motor bearing fault detection using spectral kurtosis-based feature extraction coupled with K-nearest neighbor distance analysis, IEEE Trans. Ind. Electron. 63 (3) (2015) 1793–1803.

[223] More A. Vishwendra, Pratiksha S. Salunkhe, Shivanjali V. Patil, Sumit A. Shinde, P.V. Shinde, R.G. Desavale, P.M. Jadhav, Nagaraj V. Dharwadkar, A novel method to classify rolling element bearing faults using K-nearest neighbor machine learning algorithm, ASCE- ASME J. Risk Uncertain. Eng. Syst. Part B: Mech. Eng. 8 (3) (2022) 031202.

[224] Shaul Hameed Syed, V. Muralidharan, Pradeep Kumar, Effect of K-Nearest Neighbours (k-NN) in Classifying Planetary Gearbox Faults Using Statistical Features, Technical report, SAE Technical Paper, 2022.

[225] Mohammad Zawad Ali, Md Nasmus Sakib Khan Shabbir, Xiaodong Liang, Yu Zhang, Ting Hu, Machine learning-based fault diagnosis for single-and multi-faults in induction motors using measured stator currents and vibration signals, IEEE Trans. Ind. Appl. 55 (3) (2019) 2378–2391.

[226] Firas Ben Abid, Marwen Sallem, Ahmed Braham, Optimized SWPT and decision tree for incipient bearing fault diagnosis, in: 2019 19th International Conference on Sciences and Techniques of Automatic Control and Computer Engineering, STA, IEEE, 2019, pp. 231–236.

[227] Anurag Choudhary, Deepam Goyal, Shimi Sudha Letha, Infrared thermography-based fault diagnosis of induction motor bearings using machine learning, IEEE Sensors J. 21 (2) (2020) 1727–1734.

[228] Sayanjit Singha Roy, Sayantan Dey, Soumya Chatterjee, Autocorrelation aided random forest classifier-based bearing fault detection framework, IEEE Sensors J. 20 (18) (2020) 10792–10800.

[229] P.G. Sreenath, Gopalakrishnan Praveen Kumare, Sundar Pravin, K.N. Vikram, M. Saimurugan, Automobile gearbox fault diagnosis using naive Bayes and decision tree algorithm, Appl. Mech. Mater. 813 (2015) 943–948.

[230] Mohammad Hosseinpour-Zarnaq, Mahmoud Omid, Ebrahim Biabani-Aghdam, Fault diagnosis of tractor auxiliary gearbox using vibration analysis and random forest classifier, Inf. Process. Agric. 9 (1) (2022) 60–67.

[231] Rodrigo H. Cunha Palácios, Ivan Nunes Da Silva, Alessandro Goedtel, Wagner F. Godoy, A comprehensive evaluation of intelligent classifiers for fault identification in three-phase induction motors, Electr. Power Syst. Res. 127 (2015) 249–258.

[232] Çiğdem Bakır, Different induction motor faults by new proposed random forest method, Balk. J. Electr. Comput. Eng. 11 (4) (2023) 380–386.

[233] Wenqin Zhao, Yaqiong Lv, Xiaoling Guo, Junhao Huo, An investigation on early fault diagnosis based on naive Bayes model, in: 2022 7th International Conference on Control and Robotics Engineering, ICCRE, 2022, pp. 32–36.

[234] Nannan Zhang, Lifeng Wu, Jing Yang, Yong Guan, Naive bayes bearing fault diagnosis based on enhanced independence of data, Sensors 18 (2) (2018) 463.

[235] Kiran Vernekar, Hemantha Kumar, K.V. Gangadharan, Engine gearbox fault diagnosis using empirical mode decomposition method and naïve Bayes algorithm, Sādhanā 42 (2017) 1143–1153.

[236] S. Esakimuthu Pandarakone, Santhosh Gunasekaran, Yukio Mizuno, Hisahide Nakamura, Application of naive bayes classifier theorem in detecting induction motor bearing failure, in: 2018 XIII International Conference on Electrical Machines, ICEM, IEEE, 2018, pp. 1761–1767.

[237] Shuhui Wang, Jiawei Xiang, Yongteng Zhong, Yuqing Zhou, Convolutional neural network-based hidden Markov models for rolling element bearing fault identification, Knowl.-Based Syst. 144 (2018) 65–76.

[238] Cristian López, Ángel Naranjo, Siliang Lu, Keegan J Moore, Hidden Markov model based stochastic resonance and its application to bearing fault diagnosis, J. Sound Vib. 528 (2022) 116890.

[239] Xin Li, Viliam Makis, Hongfu Zuo, Jing Cai, Optimal Bayesian control policy for gear shaft fault detection using hidden semi-Markov model, Comput. Ind. Eng. 119 (2018) 21–35.

[240] Gang Cheng, Hongyu Li, Xiao Hu, Xihui Chen, Houguang Liu, Fault diagnosis of gearbox based on local mean decomposition and discrete hidden Markov models, Proc. Inst. Mech. Eng. Part C: J. Mech. Eng. Sci. 231 (14) (2017) 2706–2717.

[241] Abdenour Soualhi, Guy Clerc, Hubert Razik, A. Lebaroud, Fault detection and diagnosis of induction motors based on hidden Markov model, in: 2012 XXth International Conference on Electrical Machines, IEEE, 2012, pp. 1693–1699.

[242] Omid Geramifard, Jian-Xin Xu, Sanjib Kumar Panda, Fault detection and diagnosis in synchronous motors using hidden Markov model-based semi-nonparametric approach, Eng. Appl. Artif. Intell. 26 (8) (2013) 1919–1929.

[243] Christos T. Yiakopoulos, Konstantinos C. Gryllias, Ioannis A. Antoniadis, Rolling element bearing fault detection in industrial environments based on a K-means clustering approach, Expert Syst. Appl. 38 (3) (2011) 2888–2911.

[244] Mir Mohammad Ettefagh, Manizheh Ghaemi, M. Yazdanian Asr, Bearing fault diagnosis using hybrid genetic algorithm K-means clustering, in: 2014 IEEE International Symposium on Innovations in Intelligent Systems and Applications (INISTA) Proceedings, IEEE, 2014, pp. 84–89.

[245] Jingbao Hou, Yunxin Wu, Abdulrahaman Shuaibu Ahmad, Hai Gong, Lei Liu, A novel rolling bearing fault diagnosis method based on adaptive feature selection and clustering, Ieee Access 9 (2021) 99756–99767.

[246] Li-Ming Wang, Yi-Min Shao, Crack fault classification for planetary gearbox based on feature selection technique and K-means clustering method, Chin. J. Mech. Eng. 31 (2018) 1–11.

[247] Imthiyas Manarikkal, Faris Elasha, David Mba, Diagnostics and prognostics of planetary gearbox using CWT, auto regression (AR) and K-means algorithm, Appl. Acoust. 184 (2021) 108314.

[248] Abdenour Soualhi, Guy Clerc, Hubert Razik, Detection and diagnosis of faults in induction motor using an improved artificial ant clustering technique, IEEE Trans. Ind. Electron. 60 (9) (2012) 4053–4062.
[249] David Camarena-Martinez, Martin Valtierra-Rodriguez, Juan P. Amezquita-Sanchez, David Granados-Lieberman, Rene J. Romero-Troncoso, Arturo Garcia-Perez, Shannon entropy and K-means method for automatic diagnosis of broken rotor bars in induction motors using vibration signals, Shock. Vib. 2016 (1) (2016) 4860309.
[250] Le Yao, Zhiqiang Ge, Cooperative deep dynamic feature extraction and variable time-delay estimation for industrial quality prediction, IEEE Trans. Ind. Informat. 17 (6) (2020) 3782–3792.
[251] Prasanna Tamilselvan, Pingfeng Wang, Failure diagnosis using deep belief learning based health state classification, Reliab. Eng. Syst. Saf. 115 (2013) 124–135.
[252] Yann LeCun, Yoshua Bengio, Geoffrey Hinton, Deep learning, Nature 521 (7553) (2015) 436–444.
[253] Tao Zhang, Feiyun Xu, Minping Jia, A centrifugal fan blade damage identification method based on the multi-level fusion of vibro-acoustic signals and CNN, Measurement 199 (2022) 111475.
[254] I-Hsi Kao, Wei-Jen Wang, Yi-Horng Lai, Jau-Woei Perng, Analysis of permanent magnet synchronous motor fault diagnosis based on learning, IEEE Trans. Instrum. Meas. 68 (2) (2018) 310–324.
[255] Li Yongbo, Du Xiaoqiang, Wan Fangyi, Wang Xianzhi, Yu Huangchao, Rotating machinery fault diagnosis based on convolutional neural network and infrared thermal imaging, Chin. J. Aeronaut. 33 (2) (2020) 427–438.
[256] Alexander Prosvirin, JaeYoung Kim, Jong-Myon Kim, Bearing fault diagnosis based on convolutional neural networks with kurtogram representation of acoustic emission signals, in: Advances in Computer Science and Ubiquitous Computing: CSA-CUTE 17, Springer, 2018, pp. 21–26.
[257] Xiaojie Guo, Liang Chen, Changqing Shen, Hierarchical adaptive deep convolution neural network and its application to bearing fault diagnosis, Measurement 93 (2016) 490–502.
[258] Wang Fuan, Jiang Hongkai, Shao Haidong, Duan Wenjing, Wu Shuaipeng, An adaptive deep convolutional neural network for rolling bearing fault diagnosis, Meas. Sci. Technol. 28 (9) (2017) 095005.
[259] Haidong Shao, Hongkai Jiang, Haizhou Zhang, Wenjing Duan, Tianchen Liang, Shuaipeng Wu, Rolling bearing fault feature learning using improved convolutional deep belief network with compressed sensing, Mech. Syst. Signal Process. 100 (2018) 743–765.
[260] Haidong Shao, Hongkai Jiang, Haizhou Zhang, Tianchen Liang, Electric locomotive bearing fault diagnosis using a novel convolutional deep belief network, IEEE Trans. Ind. Electron. 65 (3) (2017) 2727–2736.
[261] Wenfeng Gong, Hui Chen, Zehui Zhang, Meiling Zhang, Ruihan Wang, Cong Guan, Qin Wang, A novel deep learning method for intelligent fault diagnosis of rotating machinery based on improved CNN-SVM and multichannel data fusion, Sensors 19 (7) (2019) 1693.
[262] Luyang Jing, Taiyong Wang, Ming Zhao, Peng Wang, An adaptive multi-sensor data fusion method based on deep convolutional neural networks for fault diagnosis of planetary gearbox, Sensors 17 (2) (2017) 414.
[263] Huipeng Chen, Niaoqing Hu, Zhe Cheng, Lun Zhang, Yu Zhang, A deep convolutional neural network based fusion method of two-direction vibration signal data for health state identification of planetary gearboxes, Measurement 146 (2019) 268–278.
[264] Te Han, Chao Liu, Linjiang Wu, Soumik Sarkar, Dongxiang Jiang, An adaptive spatiotemporal feature learning approach for fault diagnosis in complex systems, Mech. Syst. Signal Process. 117 (2019) 170–187.
[265] Ruonan Liu, Guotao Meng, Boyuan Yang, Chuang Sun, Xuefeng Chen, Dislocated time series convolutional neural architecture: An intelligent fault diagnosis approach for electric machine, IEEE Trans. Ind. Informat. 13 (3) (2016) 1310–1320.
[266] Hong-bai Yang, Jiang-an Zhang, Lei-lei Chen, Hong-li Zhang, Shu-lin Liu, Fault diagnosis of reciprocating compressor based on convolutional neural networks with multisource raw vibration signals, Math. Probl. Eng. 2019 (2019).
[267] Min Xia, Teng Li, Lin Xu, Lizhi Liu, Clarence W De Silva, Fault diagnosis for rotating machinery using multiple sensors and convolutional neural networks, IEEE/ASME Trans. Mechatronics 23 (1) (2017) 101–110.
[268] Duy-Tang Hoang, Hee-Jun Kang, Rolling element bearing fault diagnosis using convolutional neural network and vibration image, Cogn. Syst. Res. 53 (2019) 42–50.
[269] Weiting Zhang, Dong Yang, Hongchao Wang, Xuefeng Huang, Mikael Gidlund, CarNet: A dual correlation method for health perception of rotating machinery, IEEE Sensors J. 19 (16) (2019) 7095–7106.
[270] Duy Tang Hoang, Hee Jun Kang, A motor current signal-based bearing fault diagnosis using deep learning and information fusion, IEEE Trans. Instrum. Meas. 69 (6) (2019) 3325–3333.
[271] Huaqing Wang, Shi Li, Liuyang Song, Lingli Cui, A novel convolutional neural network based fault recognition method via image fusion of multi-vibration-signals, Comput. Ind. 105 (2019) 182–190.
[272] Zhong-Xu Hu, Yan Wang, Ming-Feng Ge, Jie Liu, Data-driven fault diagnosis method based on compressed sensing and improved multiscale network, IEEE Trans. Ind. Electron. 67 (4) (2019) 3216–3225.
[273] Huaqing Wang, Shi Li, Liuyang Song, Lingli Cui, Pengxin Wang, An enhanced intelligent diagnosis method based on multi-sensor image fusion via improved deep learning network, IEEE Trans. Instrum. Meas. 69 (6) (2019) 2648–2657.
[274] Long Wen, Xinyu Li, Liang Gao, Yuyan Zhang, A new convolutional neural network-based data-driven fault diagnosis method, IEEE Trans. Ind. Electron. 65 (7) (2017) 5990–5998.
[275] Anurag Choudhary, Tauheed Mian, Shahab Fatima, Convolutional neural network based bearing fault diagnosis of rotating machine using thermal images, Measurement 176 (2021) 109196.
[276] Amin Nasiri, Amin Taheri-Garavand, Mahmoud Omid, Giovanni Maria Carlomagno, Intelligent fault diagnosis of cooling radiator based on deep learning analysis of infrared thermal images, Appl. Therm. Eng. 163 (2019) 114410.
[277] He Zhiyi, Shao Haidong, Zhong Xiang, Yang Yu, Cheng Junsheng, An intelligent fault diagnosis method for rotor-bearing system using small labeled infrared thermal images and enhanced CNN transferred from CAE, Adv. Eng. Informat. 46 (2020) 101150.
[278] Haidong Shao, Min Xia, Guangjie Han, Yu Zhang, Jiafu Wan, Intelligent fault diagnosis of rotor-bearing system under varying working conditions with modified transfer convolutional neural network and thermal images, IEEE Trans. Ind. Informat. 17 (5) (2021) 3488–3496.
[279] Dong Wang, ZhiQiang Chen, Chuan Li, René-Vinicio Sanchez, Gearbox fault identification and classification with convolutional neural networks, Shock. Vib. 2015 (2015) 390134.
[280] Yong Yao, Honglei Wang, Shaobo Li, Zhonghao Liu, Gui Gui, Yabo Dan, Jianjun Hu, End-to-end convolutional neural network model for gear fault diagnosis based on sound signals, Appl. Sci. 8 (9) (2018) 1584.
[281] Chen Lu, Zhenya Wang, Bo Zhou, Intelligent fault diagnosis of rolling bearing using hierarchical convolutional network based health state classification, Adv. Eng. Informat. 32 (2017) 139–151.
[282] Mukesh Bhadane, K.I. Ramachandran, Bearing fault identification and classification with convolutional neural network, in: 2017 International Conference on Circuit, Power and Computing Technologies, ICCPCT, IEEE, 2017, pp. 1–5.
[283] Shaobo Li, Guokai Liu, Xianghong Tang, Jianguang Lu, Jianjun Hu, An ensemble deep convolutional neural network model with improved ds evidence fusion for bearing fault diagnosis, Sensors 17 (8) (2017) 1729.
[284] Viet Tra, Sheraz Ali Khan, Jong-Myon Kim, Diagnosis of bearing defects under variable speed conditions using energy distribution maps of acoustic emission spectra and convolutional neural networks, J. Acoust. Soc. Am. 144 (4) (2018) EL322–EL327.
[285] Yong Li, Gang Cheng, Yusong Pang, Moshen Kuai, Planetary gear fault diagnosis via feature image extraction based on multi central frequencies and vibration signal frequency spectrum, Sensors 18 (6) (2018) 1735.
[286] David Verstraete, Andrés Ferrada, Enrique López Droguett, Viviana Meruane, Mohammad Modarres, et al., Deep learning enabled fault diagnosis using time-frequency image analysis of rolling element bearings, Shock. Vib. 2017 (2017).
[287] Vibhor Pandhare, Jaskaran Singh, Jay Lee, Convolutional neural network based rolling-element bearing fault diagnosis for naturally occurring and progressing defects using time-frequency domain features, in: 2019 Prognostics and System Health Management Conference, PHM-Paris, IEEE, 2019, pp. 320–326.
[288] Yu Xin, Shunming Li, Chun Cheng, Jinrui Wang, An intelligent fault diagnosis method of rotating machinery based on deep neural networks and time-frequency analysis, J. Vibroengineering 20 (6) (2018) 2321–2335.
[289] Wenxin Yu, Shoudao Huang, Weihong Xiao, Fault diagnosis based on an approach combining a spectrogram and a convolutional neural network with application to a wind turbine system, Energies 11 (10) (2018) 2561.
[290] Long Wen, Liang Gao, Xinyu Li, A new snapshot ensemble convolutional neural network for fault diagnosis, Ieee Access 7 (2019) 32037–32047.
[291] Li-Hua Wang, Xiao-Ping Zhao, Jia-Xin Wu, Yang-Yang Xie, Yong-Hong Zhang, Motor fault diagnosis based on short-time Fourier transform and convolutional neural network, Chin. J. Mech. Eng. 30 (6) (2017) 1357–1368.
[292] Xiaoxi Ding, Qingbo He, Energy-fluctuated multiscale feature learning with deep convnet for intelligent spindle bearing fault diagnosis, IEEE Trans. Instrum. Meas. 66 (8) (2017) 1926–1935.
[293] Dawei Gao, Yongsheng Zhu, Xian Wang, Ke Yan, Jun Hong, A fault diagnosis method of rolling bearing based on complex morlet CWT and CNN, in: 2018 Prognostics and System Health Management Conference, PHM-Chongqing, IEEE, 2018, pp. 1101–1105.
[294] Sheng Guo, Tao Yang, Wei Gao, Chen Zhang, Yanping Zhang, An intelligent fault diagnosis method for bearings with variable rotating speed based on pythagorean spatial pyramid pooling CNN, Sensors 18 (11) (2018) 3857.
[295] Gaowei Xu, Min Liu, Zhuofu Jiang, Dirk Söffker, Weiming Shen, Bearing fault diagnosis method based on deep convolutional neural network and random forest ensemble learning, Sensors 19 (5) (2019) 1088.
[296] M.M. Manjurul Islam, Jong-Myon Kim, Automated bearing fault diagnosis scheme using 2D representation of wavelet packet transform and deep convolutional neural network, Comput. Ind. 106 (2019) 142–153.
[297] Weifang Sun, Bin Yao, Nianyin Zeng, Binqiang Chen, Yuchao He, Xincheng Cao, Wangpeng He, An intelligent gear fault diagnosis methodology using a complex wavelet enhanced convolutional neural network, Materials 10 (7) (2017) 790.

[298] Diego Cabrera, Fernando Sancho, Chuan Li, Mariela Cerrada, René-Vinicio Sánchez, Fannia Pacheco, José Valente de Oliveira, Automatic feature extraction of time-series applied to fault severity assessment of helical gearbox in stationary and non-stationary speed operation, Appl. Soft Comput. 58 (2017) 53–64.
[299] Yan Han, Baoping Tang, Lei Deng, Multi-level wavelet packet fusion in dynamic ensemble convolutional neural network for fault diagnosis, Measurement 127 (2018) 246–255.
[300] John Grezmak, Peng Wang, Chuang Sun, Robert X. Gao, Explainable convolutional neural network for gearbox fault diagnosis, Procedia CIRP 80 (2019) 476–481.
[301] Pengfei Liang, Chao Deng, Jun Wu, Zhixin Yang, Jinxuan Zhu, Zihan Zhang, Compound fault diagnosis of gearboxes via multi-label convolutional neural network and wavelet transform, Comput. Ind. 113 (2019) 103132.
[302] Jinyang Jiao, Ming Zhao, Jing Lin, Chuancang Ding, Deep coupled dense convolutional network with complementary data for intelligent fault diagnosis, IEEE Trans. Ind. Electron. 66 (12) (2019) 9858–9867.
[303] BH Kien, D Iba, Y Ishii, Y Tsutsui, N Miura, T Iizuka, A Masuda, A Sone, I Moriwaki, Crack detection of plastic gears using a convolutional neural network pre-learned from images of meshing vibration data with transfer learning, Forsch. Im Ingenieurwesen 83 (3) (2019) 645–653.
[304] Ronny Francis Ribeiro Junior, Isac Antônio dos Santos Areias, Mateus Mendes Campos, Carlos Eduardo Teixeira, Luiz Eduardo Borges da Silva, Guilherme Ferreira Gomes, Fault detection and diagnosis in electric motors using 1d convolutional neural networks with multi-channel vibration signals, Measurement 190 (2022) 110759.
[305] Xiaoli Tang, Yuandong Xu, Xiuquan Sun, Yanfen Liu, Yu Jia, Fengshou Gu, Andrew D. Ball, Intelligent fault diagnosis of helical gearboxes with compressive sensing based non-contact measurements, ISA Trans. 133 (2023) 559–574.
[306] Guoqian Jiang, Haibo He, Jun Yan, Ping Xie, Multiscale convolutional neural networks for fault diagnosis of wind turbine gearbox, IEEE Trans. Ind. Electron. 66 (4) (2019) 3196–3207.
[307] Saisai Wang, Jie Chen, Hua Wang, Dianzhen Zhang, Degradation evaluation of slewing bearing using HMM and improved GRU, Measurement 146 (2019) 385–395.
[308] Han Liu, Jianzhong Zhou, Yang Zheng, Wei Jiang, Yuncheng Zhang, Fault diagnosis of rolling bearings with recurrent neural network-based autoencoders, ISA Trans. 77 (2018) 167–178.
[309] Guo-Ping Liao, Wei Gao, Geng-Jie Yang, Mou-Fa Guo, Hydroelectric generating unit fault diagnosis using 1-D convolutional neural network and gated recurrent unit in small hydro, IEEE Sensors J. 19 (20) (2019) 9352–9363.
[310] Honghu Pan, Xingxi He, Sai Tang, Fanming Meng, An improved bearing fault diagnosis method using one-dimensional CNN and LSTM, Strojni v Ski. Vestnik - J. Mech. Eng. 64 (7–8) (2018) 443–452.
[311] Abhinav Saxena, Rajat Kumar, Arun Kumar Rawat, Mohd Majid, Jay Singh, S. Devakirubakaran, Gyanendra Kumar Singh, et al., Abnormal health monitoring and assessment of a three-phase induction motor using a supervised CNN-RNN-based machine learning algorithm, Math. Probl. Eng. 2023 (2023).
[312] L.Y. Imamura, S.L. Avila, F.S. Pacheco, M.B.C. Salles, L.S Jablon, Diagnosis of unbalance in lightweight rotating machines using a recurrent neural network suitable for an edge-computing framework, J. Control. Autom. Electr. Syst. 33 (4) (2022) 1272–1285.
[313] Shentong Ni, Yang Tang, Xia Fang, Dingcheng Zhang, Jie Wang, Jiawei Liao, Recurrent neural network-based fault diagnosis method for multisource information fusion of drilling pump power end gear, in: 2023 CAA Symposium on Fault Detection, Supervision and Safety for Technical Processes, SAFEPROCESS, 2023, pp. 1–6.
[314] Diego Cabrera, Adriana Guamán, Shaohui Zhang, Mariela Cerrada, René-Vinicio Sánchez, Juan Cevallos, Jianyu Long, Chuan Li, Bayesian approach and time series dimensionality reduction to LSTM-based model-building for fault diagnosis of a reciprocating compressor, Neurocomputing 380 (2020) 51–66.
[315] Junchuan Shi, Dikang Peng, Zhongxiao Peng, Ziyang Zhang, Kai Goebel, Dazhong Wu, Planetary gearbox fault diagnosis using bidirectional-convolutional LSTM networks, Mech. Syst. Signal Process. 162 (2022) 107996.
[316] Dikang Peng, Wade A. Smith, Pietro Borghesani, Robert B. Randall, Zhongxiao Peng, Comprehensive planet gear diagnostics: Use of transmission error and mesh phasing to distinguish localised fault types and identify faulty gears, Mech. Syst. Signal Process. 127 (2019) 531–550.
[317] L. Zou, K.J. Zhuang, A. Zhou, J. Hu, Bayesian optimization and channel-fusion-based convolutional autoencoder network for fault diagnosis of rotating machinery, Eng. Struct. 280 (2023) 115708.
[318] Ning Jia, Yao Cheng, Yunyang Liu, Youyuan Tian, Intelligent fault diagnosis of rotating machines based on wavelet time-frequency diagram and optimized stacked denoising auto-encoder, IEEE Sensors J. 22 (17) (2022) 17139–17150.
[319] Wenan Cai, Qianqian Zhang, Jie Cui, A novel fault diagnosis method for denoising autoencoder assisted by digital twin, Comput. Intell. Neurosci. 2022 (1) (2022) 5077134.
[320] Martin Hemmer, Andreas Klausen, Huynh Van Khang, Kjell G. Robbersmyr, Tor I. Waag, Health indicator for low-speed axial bearings using variational autoencoders, IEEE Access 8 (2020) 35842–35852.
[321] Gabriel San Martin, Enrique Lopez Droguett, Viviane Meruane, Márcio das Chagas Moura, Deep variational auto-encoders: A promising tool for dimensionality reduction and ball bearing elements fault diagnosis, Struct. Heal. Monit. 18 (4) (2019) 1092–1128.
[322] Xin Wang, Hongkai Jiang, Zhenghong Wu, Qiao Yang, Adaptive variational autoencoding generative adversarial networks for rolling bearing fault diagnosis, Adv. Eng. Informat. 56 (2023) 102027.
[323] Penghao Pan, Dong Zhao, Yueyang Li, A fault diagnosis framework for rotating machinery of marine equipment: A semi-supervised learning framework based on contractive stacked autoencoder, Proc. Inst. Mech. Eng. Part M: J. Eng. Marit. Environ. 237 (3) (2023) 625–636.
[324] Haiquan Song, Wengang Ma, Zhonghe Han, Xiaoxun Zhu, Diagnosis of unbalanced rolling bearing fault sample based on adaptive sparse contrative auto-encoder and IGWO-USELM, Measurement 198 (2022) 111353.
[325] Yuyan Zhang, Liang Gao, Xiaoyu Wen, Haoqi Wang, Intelligent fault diagnosis of machine under noisy environment using ensemble orthogonal contractive auto-encoder, Expert Syst. Appl. 203 (2022) 117408.
[326] Yuyan Zhang, Xinyu Li, Liang Gao, Wen Chen, Peigen Li, Ensemble deep contractive auto-encoders for intelligent fault diagnosis of machines under noisy environment, Knowl.-Based Syst. 196 (2020) 105764.
[327] Daoguang Yang, Hamid Reza Karimi, Kangkang Sun, Residual wide-kernel deep convolutional auto-encoder for intelligent rotating machinery fault diagnosis with limited samples, Neural Netw. 141 (2021) 133–144.
[328] Quan Qian, Yi Qin, Yi Wang, Fuqiang Liu, A new deep transfer learning network based on convolutional auto-encoder for mechanical fault diagnosis, Measurement 178 (2021) 109352.
[329] Xingchen Liu, Qicai Zhou, Jiong Zhao, Hehong Shen, Xiaolei Xiong, Fault diagnosis of rotating machinery under noisy environment conditions based on a 1-D convolutional autoencoder and 1-D convolutional neural network, Sensors 19 (4) (2019) 972.
[330] Zenghui An, Xingxing Jiang, Jie Liu, Mode-decoupling auto-encoder for machinery fault diagnosis under unknown working conditions, IEEE Trans. Ind. Informat. (2023).
[331] Wennian Yu, I.I. Yong Kim, Chris Mechefske, An improved similarity-based prognostic algorithm for RUL estimation using an RNN autoencoder scheme, Reliab. Eng. Syst. Saf. 199 (2020) 106926.
[332] Zhiyuan Li, Yu Sun, Laihao Yang, Zhibin Zhao, Xuefeng Chen, Unsupervised machine anomaly detection using autoencoder and temporal convolutional network, IEEE Trans. Instrum. Meas. 71 (2022) 1–13.
[333] Geoffrey E. Hinton, Simon Osindero, Yee-Whye Teh, A Fast Learning Algorithm for Deep Belief Nets, Neural Comput. 18 (7) (2006) 1527–1554.
[334] Jingbo Gai, Kunyu Zhong, Xuejiao Du, Ke Yan, Junxian Shen, Detection of gear fault severity based on parameter-optimized deep belief network using sparrow search algorithm, Measurement 185 (2021) 110079.
[335] Yi Qin, Xin Wang, Jingqiang Zou, The optimized deep belief networks with improved logistic sigmoid units and their application in fault diagnosis for planetary gearboxes of wind turbines, IEEE Trans. Ind. Electron. 66 (5) (2019) 3814–3824.
[336] Jie Tao, Yilun Liu, Dalian Yang, et al., Bearing fault diagnosis based on deep belief network and multisensor information fusion, Shock. Vib. 2016 (2016).
[337] Zhiqiang Chen, Shengcai Deng, Xudong Chen, Chuan Li, René-Vinicio Sanchez, Huafeng Qin, Deep neural networks-based rolling bearing fault diagnosis, Microelectron. Reliab. 75 (2017) 327–333.
[338] Jialin Li, Xueyi Li, David He, Yongzhi Qu, Unsupervised rotating machinery fault diagnosis method based on integrated SAE–DBN and a binary processor, J. Intell. Manuf. 31 (2020) 1899–1916.
[339] I.J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, Y. Bengio, Generative adversarial networks, 1–9, 2014, arXiv preprint arXiv:1406.2661.
[340] Thomas Schlegl, Philipp Seeböck, Sebastian M. Waldstein, Ursula Schmidt-Erfurth, Georg Langs, Unsupervised anomaly detection with generative adversarial networks to guide marker discovery, in: Marc Niethammer, Martin Styner, Stephen Aylward, Hongtu Zhu, Ipek Oguz, Pew-Thian Yap, Dinggang Shen (Eds.), Information Processing in Medical Imaging, Springer International Publishing, Cham, 2017, pp. 146–157.
[341] Jingyao Wu, Zhibin Zhao, Chuang Sun, Ruqiang Yan, Xuefeng Chen, Ss-InfoGAN for class-imbalance classification of bearing faults, Procedia Manuf. 49 (2020) 99–104.
[342] Wei Li, Xiang Zhong, Haidong Shao, Baoping Cai, Xingkai Yang, Multi-mode data augmentation and fault diagnosis of rotating machinery using modified ACGAN designed with new framework, Adv. Eng. Informat. 52 (2022) 101552.
[343] Ming Zhang, Duo Wang, Weining Lu, Jun Yang, Zhiheng Li, Bin Liang, A deep transfer model with wasserstein distance guided multi-adversarial networks for bearing fault diagnosis under different working conditions, IEEE Access 7 (2019) 65303–65318.
[344] Xudong Mao, Qing Li, Haoran Xie, Raymond YK Lau, Zhen Wang, Stephen Paul Smolley, Least squares generative adversarial networks, in: Proceedings of the IEEE International Conference on Computer Vision, 2017, pp. 2794–2802.
[345] Junbo Zhao, Michael Mathieu, Yann LeCun, Energy-based generative adversarial network, 2016, arXiv preprint arXiv:1609.03126.

[346] David Berthelot, Thomas Schumm, Luke Metz, Began: Boundary equilibrium generative adversarial networks, 2017, arXiv preprint arXiv:1703.10717.
[347] Diwang Ruan, Xuran Chen, Clemens Gühmann, Jianping Yan, Improvement of generative adversarial network and its application in bearing fault diagnosis: A review, Lubricants 11 (2) (2023).
[348] Qingwen Guo, Yibin Li, Yanjun Liu, Shengyao Gao, Yan Song, Data augmentation for intelligent mechanical fault diagnosis based on local shared multiple-generator GAN, IEEE Sensors J. 22 (10) (2022) 9598–9609.
[349] Liang Ma, Yu Ding, Zili Wang, Chao Wang, Jian Ma, Chen Lu, An interpretable data augmentation scheme for machine fault diagnosis based on a sparsity-constrained generative adversarial network, Expert Syst. Appl. 182 (2021) 115234.
[350] Dongdong Wang, Yining Dong, Han Wang, Gang Tang, Limited fault data augmentation with compressed sensing for bearing fault diagnosis, IEEE Sensors J. (2023).
[351] Quan Sun, Fei Peng, Xianghai Yu, Hongsheng Li, Data augmentation strategy for power inverter fault diagnosis based on wasserstein distance and auxiliary classification generative adversarial network, Reliab. Eng. Syst. Saf. 237 (2023) 109360.
[352] Qianjun Liu, Guijun Ma, Cheng Cheng, Data fusion generative adversarial network for multi-class imbalanced fault diagnosis of rotating machinery, IEEE Access 8 (2020) 70111–70124.
[353] Taisheng Zheng, Lei Song, Jianxing Wang, Wei Teng, Xiaoli Xu, Chao Ma, Data synthesis using dual discriminator conditional generative adversarial networks for imbalanced fault diagnosis of rolling bearings, Measurement 158 (2020) 107741.
[354] Bingxi Zhao, Qi Yuan, Improved generative adversarial network for vibration-based fault diagnosis with imbalanced data, Measurement 169 (2021) 108522.
[355] Yi Qin, Hongyu Liu, Yongfang Mao, Faulty rolling bearing digital twin model and its application in fault diagnosis with imbalanced samples, Adv. Eng. Informat. 61 (2024) 102513.
[356] Yaoxiang Yu, Liang Guo, Hongli Gao, Yuekai Liu, PCWGAN-GP: A new method for imbalanced fault diagnosis of machines, IEEE Trans. Instrum. Meas. 71 (2022) 1–11.
[357] Xi Gu, Yaoxiang Yu, Liang Guo, Hongli Gao, Ming Luo, CSWGAN-GP: A new method for bearing fault diagnosis under imbalanced condition, Measurement 217 (2023) 113014.
[358] Wenqian Jiang, Yang Hong, Beitong Zhou, Xin He, Cheng Cheng, A GAN-based anomaly detection approach for imbalanced industrial time series, IEEE Access 7 (2019) 143608–143619.
[359] Ke Yan, Chiller fault detection and diagnosis with anomaly detective generative adversarial network, Build. Environ. 201 (2021) 107982.
[360] Jiangji Peng, Haidong Shao, Yiming Xiao, Baoping Cai, Bin Liu, Industrial surface defect detection and localization using multi-scale information focusing and enhancement ganomaly, Expert Syst. Appl. 238 (2024) 122361.
[361] Xin Xu, Hongbo Zhao, Haoqiang Liu, Hua Sun, LSTM-gan-xgboost based anomaly detection algorithm for time series data, in: 2020 11th International Conference on Prognostics and System Health Management (PHM-2020 Jinan), IEEE, 2020, pp. 334–339.
[362] Hao Wen, Djordy Van Maele, Jean Carlos Poletto, Patrick De Baets, Konstantinos Gryllias, A gear health indicator based on f-anogan, in: PHM Society European Conference, Vol. 8, No. 1, 2024, p. 12.
[363] Rui Nian, Jinfeng Liu, Biao Huang, A review on reinforcement learning: Introduction and applications in industrial process control, Comput. Chem. Eng. 139 (2020) 106886.
[364] Zhi-Hua Zhou, A brief introduction to weakly supervised learning, Natl. Sci. Rev. 5 (1) (2018) 44–53.
[365] R.S. Sutton, A.G. Barto, Reinforcement learning: An introduction, 2018.
[366] Hamid Teimourzadeh, Arash Moradzadeh, Maryam Shoaran, Behnam Mohammadi-Ivatloo, Reza Razzaghi, High impedance single-phase faults diagnosis in transmission lines via deep reinforcement learning of transfer functions, IEEE Access 9 (2021) 15796–15809.
[367] Asma Belhadi, Youcef Djenouri, Gautam Srivastava, Alireza Jolfaei, Jerry Chun-Wei Lin, Privacy reinforcement learning for faults detection in the smart grid, Ad Hoc Netw. 119 (2021) 102541.
[368] Li Junhuai, Wu Yunwen, Wang Huaijun, Xu Jiang, Fault detection method based on adversarial reinforcement learning, Front. Comput. Sci. 4 (2023) 1007665.
[369] Jifei Deng, Seppo Sierla, Jie Sun, Valeriy Vyatkin, Offline reinforcement learning for industrial process control: A case study from steel industry, Inform. Sci. 632 (2023) 221–231.
[370] Tony Z. Zhao, Jianlan Luo, Oleg Sushkov, Rugile Pevceviciute, Nicolas Heess, Jon Scholz, Stefan Schaal, Sergey Levine, Offline meta-reinforcement learning for industrial insertion, in: 2022 International Conference on Robotics and Automation, ICRA, IEEE, 2022, pp. 6386–6393.
[371] Sergey Levine, Aviral Kumar, George Tucker, Justin Fu, Offline reinforcement learning: Tutorial, review, and perspectives on open problems, 2020, arXiv preprint arXiv:2005.01643.
[372] Cheng Qian, Wei Yu, Xing Liu, David Griffith, Nada Golmie, Towards online continuous reinforcement learning on industrial internet of things, in: 2021 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/IOP/SCI), IEEE, 2021, pp. 280–287.
[373] Siyu Zhang, S.U. Lei, G.U. Jiefei, L.I. Ke, Zhou Lang, Michael Pecht, Rotating machinery fault detection and diagnosis based on deep domain adaptation: A survey, Chin. J. Aeronaut. 36 (1) (2023) 45–74.
[374] Hanting Zhou, Wenhe Chen, Longsheng Cheng, Darren Williams, Clarence W. De Silva, Min Xia, Reliable and intelligent fault diagnosis with evidential VGG neural networks, IEEE Trans. Instrum. Meas. 72 (2023) 1–12.
[375] Changbo He, Yujie Cao, Yang Yang, Yongbin Liu, Xianzeng Liu, Zheng Cao, Fault diagnosis of rotating machinery based on the improved multidimensional normalization ResNet, IEEE Trans. Instrum. Meas. (2023).
[376] Pengfei Liang, Wenhui Wang, Xiaoming Yuan, Siyuan Liu, Lijie Zhang, Yiwei Cheng, Intelligent fault diagnosis of rolling bearing based on wavelet transform and improved ResNet under noisy labels and environment, Eng. Appl. Artif. Intell. 115 (2022) 105269.
[377] Xiaohan Chen, Rui Yang, Yihao Xue, Mengjie Huang, Roberto Ferrero, Zidong Wang, Deep transfer learning for bearing fault diagnosis: A systematic review since 2016, IEEE Trans. Instrum. Meas. 72 (2023) 1–21.
[378] Xingqiu Li, Hongkai Jiang, Shaowei Liu, Jianjun Zhang, Jun Xu, A unified framework incorporating predictive generative denoising autoencoder and deep coral network for rolling bearing fault diagnosis with unbalanced data, Measurement 178 (2021) 109345.
[379] Xingkai Chen, Haidong Shao, Yiming Xiao, Shen Yan, Baoping Cai, Bin Liu, Collaborative fault diagnosis of rotating machinery via dual adversarial guided unsupervised multi-domain adaptation network, Mech. Syst. Signal Process. 198 (2023) 110427.
[380] Lanjun Wan, Yuanyuan Li, Keyu Chen, Kun Gong, Changyun Li, A novel deep convolution multi-adversarial domain adaptation model for rolling bearing fault diagnosis, Measurement 191 (2022) 110752.
[381] Yongwen Tan, Liang Guo, Hongli Gao, Zhibin Lin, Yuekai Liu, MiDAN: A framework for cross-domain intelligent fault diagnosis with imbalanced datasets, Measurement 183 (2021) 109834.
[382] Liang Guo, Yaoxiang Yu, Yuekai Liu, Hongli Gao, Tao Chen, Reconstruction domain adaptation transfer network for partial transfer learning of machinery fault diagnostics, IEEE Trans. Instrum. Meas. 71 (2021) 1–10.
[383] Quan Qian, Jun Luo, Yi Qin, Adaptive intermediate class-wise distribution alignment: a universal domain adaptation and generalization method for machine fault diagnosis, IEEE Trans. Neural Netw. Learn. Syst. (2024).
[384] Quan Qian, Huayan Pu, Tianjia Tu, Yi Qin, Variance discrepancy representation: A vibration characteristic-guided distribution alignment metric for fault transfer diagnosis, Mech. Syst. Signal Process. 217 (2024) 111544.
[385] Huimin Zhao, Xiaoxu Yang, Baojie Chen, Huayue Chen, Wu Deng, Bearing fault diagnosis using transfer learning and optimized deep belief network, Meas. Sci. Technol. 33 (6) (2022) 065009.
[386] Zitong Wan, Rui Yang, Mengjie Huang, Deep transfer learning-based fault diagnosis for gearbox under complex working conditions, Shock. Vib. 2020 (2020) 1–13.
[387] Zhongwei Zhang, Huaihai Chen, Shunming Li, Zenghui An, Jinrui Wang, A novel geodesic flow kernel based domain adaptation approach for intelligent fault diagnosis under varying working condition, Neurocomputing 376 (2020) 54–64.
[388] Faras Jamil, Timothy Verstraeten, Ann Nowé, Cédric Peeters, Jan Helsen, A deep boosted transfer learning method for wind turbine gearbox fault detection, Renew. Energy 197 (2022) 331–341.
[389] Hongyu Zhong, Yong Lv, Rui Yuan, Di Yang, Bearing fault diagnosis using transfer learning and self-attention ensemble lightweight convolutional neural network, Neurocomputing 501 (2022) 765–777.
[390] Guansong Pang, Chunhua Shen, Longbing Cao, Anton Van Den Hengel, Deep learning for anomaly detection: A review, ACM Comput. Surv. 54 (2) (2021) 1–38.
[391] Shen Yan, Haidong Shao, Zhishan Min, Jiangji Peng, Baoping Cai, Bin Liu, FGDAE: A new machinery anomaly detection method toward complex operating conditions, Reliab. Eng. Syst. Saf. 236 (2023) 109319.
[392] Christian Velasco-Gallego, Iraklis Lazakis, RADIS: A real-time anomaly detection intelligent system for fault diagnosis of marine machinery, Expert Syst. Appl. 204 (2022) 117634.
[393] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, Illia Polosukhin, Attention is all you need, Adv. Neural Inf. Process. Syst. 30 (2017).
[394] Haiyue Wu, Matthew J. Triebe, John W. Sutherland, A transformer-based approach for novel fault detection and fault classification/diagnosis in manufacturing: A rotary system application, J. Manuf. Syst. 67 (2023) 439–452.
[395] Zhiying Wang, Zheng Zhou, Wengang Xu, Chuang Sun, Ruqiang Yan, Physics informed neural networks for fault severity identification of axial piston pumps, J. Manuf. Syst. 71 (2023) 421–437.

[396] Yigit A. Yucesan, Felipe A.C. Viana, A physics-informed neural network for wind turbine main bearing fatigue, Int. J. Progn. Heal. Manag. 11 (1) (2020).
[397] Qing Ni, J.C. Ji, Benjamin Halkon, Ke Feng, Asoke K. Nandi, Physics-informed residual network (piResNet) for rolling element bearing fault diagnostics, Mech. Syst. Signal Process. 200 (2023) 110544.
[398] Aymen Rayane Khouas, Mohamed Reda Bouadjenek, Hakim Hacid, Sunil Aryal, Training machine learning models at the edge: A survey, 2024, arXiv preprint arXiv:2403.02619.
[399] Manan Mehta, Siyuan Chen, Haichuan Tang, Chenhui Shao, A federated learning approach to mixed fault diagnosis in rotating machinery, J. Manuf. Syst. 68 (2023) 687–694.
[400] Funa Zhou, Shun Liu, Hamido Fujita, Xiong Hu, Yanbo Zhang, Bei Wang, Kai Wang, Fault diagnosis based on federated learning driven by dynamic expansion for model layers of imbalanced client, Expert Syst. Appl. 238 (2024) 121982.
[401] Tao Wen, Xiaohan Chen, Dingcheng Zhang, Clive Roberts, Baigen Cai, A sequential and asynchronous federated learning framework for railway point machine fault diagnosis with imperfect data transmission, IEEE Trans. Ind. Informat. 20 (6) (2024) 8828–8837.
[402] Bin Yang, Yaguo Lei, Xiang Li, Naipeng Li, Targeted transfer learning through distribution barycenter medium for intelligent fault diagnosis of machines with data decentralization, Expert Syst. Appl. 244 (2024) 122997.
[403] Wei Zhang, Ziwei Wang, Xiang Li, Blockchain-based decentralized federated transfer learning methodology for collaborative machinery fault diagnosis, Reliab. Eng. Syst. Saf. 229 (2023) 108885.
[404] Quan Qian, Jun Luo, Yi Qin, Heterogeneous federated domain generalization network with common representation learning for cross-load machinery fault diagnosis, IEEE Trans. Syst. Man Cybern. Syst. (2024).
[405] Yuyan Zhang, Xinyu Li, Liang Gao, Lihui Wang, Long Wen, Imbalanced data fault diagnosis of rotating machinery using synthetic oversampling and feature learning, J. Manuf. Syst. 48 (2018) 34–50, Special Issue on Smart Manufacturing.
[406] Kiran Maharana, Surajit Mondal, Bhushankumar Nemade, A review: Data pre-processing and data augmentation techniques, Glob. Transitions Proc. 3 (1) (2022) 91–99, International Conference on Intelligent Engineering Approach(ICIEA-2022).
[407] Rakesh A. Patel, Bhavesh R. Bhalja, Condition monitoring and fault diagnosis of induction motor using support vector machine, Electr. Power Components Syst. 44 (6) (2016) 683–692.
[408] Dhiraj Neupane, Mohamed Reda Bouadjenek, Richard Dazeley, Sunil Aryal, Machinery fault detection using advanced machine learning techniques, in: PHM Society European Conference, Vol. 8, No. 1, 2024, p. 4.
[409] Francois Chollet, The limitations of deep learning, in: Deep Learning with Python, first ed., Manning Publications, New York, NY, 2017.
[410] Daniel Camilleri, Tony Prescott, Analysing the limitations of deep learning for developmental robotics, in: Michael Mangan, Mark Cutkosky, Anna Mura, Paul F.M.J. Verschure, Tony Prescott, Nathan Lepora (Eds.), Biomimetic and Biohybrid Systems, Springer International Publishing, Cham, 2017, pp. 86–94.
[411] Laith Alzubaidi, Jinglan Zhang, Amjad J. Humaidi, Ayad Al-Dujaili, Ye Duan, Omran Al-Shamma, J. Santamaría, Mohammed A. Fadhel, Muthana Al-Amidie, Laith Farhan, Review of deep learning: concepts, CNN architectures, challenges, applications, future directions, J. Big Data 8 (1) (2021) 53.
[412] Purushottam Gangsar, Aditya Raj Bajpei, Rajkumar Porwal, A review on deep learning based condition monitoring and fault diagnosis of rotating machinery, Noise Vib. Worldw. 53 (11) (2022) 550–578.
[413] Andrew Ng, What artificial intelligence can and can't do right now, Harv. Bus. Rev. (2016).
[414] Qingchen Zhang, Laurence T. Yang, Zhikui Chen, Peng Li, A survey on deep learning for big data, Inf. Fusion 42 (2018) 146–157.
[415] Syahril Ramadhan Saufi, Zair Asrar Bin Ahmad, Mohd Salman Leong, Meng Hee Lim, Challenges and opportunities of deep learning models for machinery fault detection and diagnosis: A review, IEEE Access 7 (2019) 122644–122662.
[416] Diego Andina, Athanasios Voulodimos, Nikolaos Doulamis, Anastasios Doulamis, Eftychios Protopapadakis, Deep learning for computer vision: A brief review, Comput. Intell. Neurosci. 2018 (2018) 7068349.
[417] Divya Saxena, Jiannong Cao, Generative adversarial networks (GANs): Challenges, solutions, and future directions, ACM Comput. Surv. 54 (3) (2021).
[418] Mohanad Nawfal Mustafa, Classification of maintenance techniques and diagnosing failures methods, J. Phys. Conf. Ser. 2060 (1) (2021) 012014.
[419] W. Wang, K. Vos, J. Taylor, C. Jenkins, B. Bala, L. Whitehead, Z. Peng, Is deep learning superior to traditional techniques in machine health monitoring applications, Aeronaut. J. (2023) 1–13.
[420] Yaoxiang Yu, Liang Guo, Hongli Gao, Yuekai Liu, Tingting Feng, Pareto-optimal adaptive loss residual shrinkage network for imbalanced fault diagnostics of machines, IEEE Trans. Ind. Informat. 18 (4) (2022) 2233–2243.
[421] Hasmat Malik, Yogesh Pandya, Aakash Parashar, Rajneesh Sharma, Feature extraction using EMD and classifier through artificial neural networks for gearbox fault diagnosis, in: Hasmat Malik, Smriti Srivastava, Yog Raj Sood, Aamir Ahmad (Eds.), Applications of Artificial Intelligence Techniques in Engineering, Springer Singapore, Singapore, 2019, pp. 309–317.
[422] Wenyi Wang, David Blunt, Joanna Kappas, A description of the dataset for HUMS2023 data challenge, 2022.
[423] José Miguel Ramírez-Sanz, Jose-Alberto Maestro-Prieto, Álvar Arnaiz-González, Andrés Bustillo, Semi-supervised learning for industrial fault detection and diagnosis: A systemic review, ISA Trans. (2023).
[424] Zhenxiang Li, CWRU bearing dataset and gearbox dataset of IEEE PHM challenge competition in 2009, 2019.
[425] Xinghui Li, 2010 PHM society conference data challenge, 2021.
[426] Pei Cao, Shengli Zhang, Jiong Tang, Preprocessing-free gear fault diagnosis using small datasets with deep convolutional neural network-based transfer learning, IEEE Access 6 (2018) 26241–26253.
[427] Minglei Zheng, Qi Chang, Junfeng Man, Yi Liu, Yiping Shen, Two-stage multi-scale fault diagnosis method for rolling bearings with imbalanced data, Machines 10 (5) (2022).

Mr Dhiraj Neupane is a Deakin University PhD candidate specializing in intelligent fault detection using advanced machine learning algorithms. He holds a master's degree in information and communication engineering from Changwon National University, South Korea. His research focuses on machinery fault detection and signal processing, with additional interests in computer vision and sonar signal processing, leading to several high-impact publications. Mr. Neupane has also actively contributed as a reviewer for various prestigious journals.

Dr Mohamed Reda Bouadjenek is a Senior Lecturer (Assistant Professor) of Applied Artificial Intelligence within the School of Information Technology at Deakin University, Australia. His previous roles include Research Fellow positions at The University of Toronto (2017-2019) and The University of Melbourne (2015-2017). Prior to that, he served as a postdoc researcher at the French Institute for Research in Computer Science and Automation (INRIA) from 2014 to 2015 in France. Reda holds a Ph.D. and an MSc in Computer Science from the University of ParisSaclay, France, earned in 2013 and 2009, respectively. He also obtained a BSc in Computer Science from the University of Science and Technology Houari Boumediene in 2008. Reda's research expertise encompasses a wide range of topics within the data-driven domains of Machine Learning, Deep Learning, and Information Retrieval. Leveraging analytical and algorithmic tools from these fields, he addresses real-world challenges across various applications, including recommender systems, interactive visual search interfaces, social network analysis, and data quality.

Prof. Richard Dazeley is the Deputy Head of School for the School of Information Technology (SIT) at Deakin University. He is an internationally leading researcher in Multiobjective Reinforcement Learning (MORL) and its application to the Human-alignment problem. He is a senior member of the International AI Existential Safety Community with the Future of Life Institute; founding and leading member of the International Multiobjective Reinforcement Learning community; and, the co-founder/leader of the Australian Responsible Autonomous Agents Collective (ARAAC). As one of the founding researchers of MORL he is an author of several of the key papers in the field; frequent internationally invited speaker; guest editor of journals; and workshop organizer in the field.

A/Prof. Sunil Aryal is an associate professor in Data Science at the School of IT at Deakin University Australia, where he co-leads the Machine Learning for Decision Support (MLDS) Research Group. His research interests are in the areas of Artificial Intelligence (AI), Machine Learning (ML) and Data Mining (DM). He is working on anomaly detection, clustering, kernel/similarity-based learning, ensemble methods, learning from heterogeneous, noisy and limited/small data, reinforcement learning, natural language processing, and computer vision. He is particularly interested in applying AI/ML/DM models to solve real-world problems in Defence, Cybersecurity, Engineering, Agriculture, Healthcare and Education. He has published over 75 papers in top-tier international venues and secured over AUD 4.5M in external research funding from agencies like the US and Australian Defence, State Department of Education and Training Victoria, Technology Innovation Institute UAE, and Table Tennis Australia.