algorithms

*Review*

# A Survey on Fault Diagnosis of Rolling Bearings

**Bo Peng [1], Ying Bi [2,3,*], Bing Xue [3], Mengjie Zhang [3] and Shuting Wan [4]**

1 College of Mechanical and Electrical Engineering, Hebei Agricultural University, Baoding 071000, China
2 School of Electrical and Information Engineering, Zhengzhou University, Zhengzhou 450001, China
3 School of Engineering and Computer Science, Victoria University of Wellington, Wellington 6140, New Zealand
4 Hebei Key Laboratory of Electric Machinery Health Maintenance & Failure Prevention, North China Electric Power University, Baoding 071003, China
* Correspondence: ying.bi@ecs.vuw.ac.nz

**Abstract:** The failure of a rolling bearing may cause the shutdown of mechanical equipment and even induce catastrophic accidents, resulting in tremendous economic losses and a severely negative impact on society. Fault diagnosis of rolling bearings becomes an important topic with much attention from researchers and industrial pioneers. There are an increasing number of publications on this topic. However, there is a lack of a comprehensive survey of existing works from the perspectives of fault detection and fault type recognition in rolling bearings using vibration signals. Therefore, this paper reviews recent fault detection and fault type recognition methods using vibration signals. First, it provides an overview of fault diagnosis of rolling bearings and typical fault types. Then, existing fault diagnosis methods are categorized into fault detection methods and fault type recognition methods, which are separately revised and discussed. Finally, a summary of existing datasets, limitations/challenges of existing methods, and future directions are presented to provide more guidance for researchers who are interested in this field. Overall, this survey paper conducts a review and analysis of the methods used to diagnose rolling bearing faults and provide comprehensive guidance for researchers in this field.

**Keywords:** rolling bearing; diagnosis; fault detection; fault type recognition; signal processing; machine learning



**Citation:** Peng, B.; Bi, Y.; Xue, B.; Zhang, M.; Wan, S. A Survey on Fault Diagnosis of Rolling Bearings. *Algorithms* **2022**, *15*, 347. https://doi.org/10.3390/a15100347

Academic Editor: Sergio Rajsbaum

Received: 22 August 2022
Accepted: 22 September 2022
Published: 26 September 2022



## 1. Introduction

With the rapid development of technology and science, modern industry has become increasingly important in our daily life. The advancement of science and technology has led to the gradual development of large-scale and high-speed rotating machinery with integration, precision, and intelligence. Rotating machinery is an essential part of modern industry and is widely used in many fields, including energy and power, machinery manufacturing, transportation, and aerospace. Once mechanical equipment is successfully developed for production, the reliability and safety of the equipment become increasingly crucial, and the fault diagnosis and condition monitoring of the core components become an arduous task [1–3].

Roller bearings are widely used in rotating machinery and are an indispensable component that supports the rotating shaft and serves as a connector between stationary and rotating parts. Although rolling bearing damage occurs at the component level, it frequently leads to more severe equipment failures. According to statistics, rolling bearing failures account for 40–90% of all rotating machinery failures [4]. The initial failure of the rolling bearing of a wind turbine will only affect itself, and the unit will remain operational. However, as the times of abnormal operations increase, external excitations caused by broken bearings will cause the traditional system to malfunction, resulting in a fire in extreme cases. Roll bearing failure in the rolling mill will cause a reduction in the quality of rolled products, which will lead to the production line being stopped and result in significant economic losses. Due to the complex and changing conditions operating in rotating machinery, rolling bearings often fail before their designed life ends, and their actual service life is often shorter than their design life, so a routine shutdown inspection is not the best way. Therefore, an effective and intelligent fault diagnosis of rolling bearings is of considerable practical significance for ensuring the health of rotating equipment and machinery.

Fault diagnosis of rolling bearings is a multidisciplinary field that incorporates computer science, mathematics, electronics, signal processing, engineering, and other modern technologies. Rolling bearing fault diagnosis is to diagnose the bearing health status through the collected operation data. Fault diagnosis can be broadly categorized into fault detection and fault type recognition. Fault detection is to detect faults from the collected data, while fault type recognition is to recognize faults and their types from the data. During the past ten years, fault diagnosis of rolling bearings has attracted considerable attention from both academics and the industry. Figure 1 shows the number of publications on the topic of rolling bearing fault diagnosis extracted from the Scopus database. It is clear that the number of publications has gradually increased from 2011 to 2021. There are several survey papers on fault diagnosis. However, most of them focus on specifical tasks or methods, such as machine learning-based methods [5] for prognostics and health management of rolling element bearings [6], Fourier transform and enhanced fast Fourier transform algorithms [7], artificial intelligence methods [8], spectral kurtosis [9], and signal processing techniques [10]. Very few of them provide a general and comprehensive survey on rolling bearing fault diagnosis using vibration signals from the perspectives of fault detection and fault type recognition.

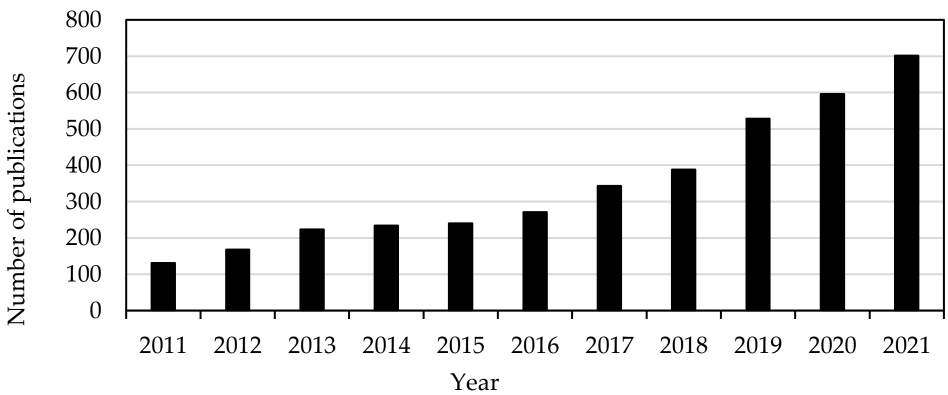



**Figure 1.** The number of publications on rolling bearing fault diagnosis from 2011 to 2021.

To address the above limitations, this paper reviews over 150 related publications in recent years, including over 100 publications from 2016 to 2021. These publications are well-known or representative ones in the rolling bearing fault diagnosis community. This survey discusses not only traditional methods based on signal processing and analysis but also machine learning and artificial intelligence methods, including feature extraction/reduction methods, deep learning methods, and evolutionary learning methods, to present a relatively full picture of this field. In addition, this survey summarizes commonly used datasets, existing limitations/challenges, and future research trends to provide researchers with useful guidance.

The structure of the survey is task-based, including tasks for fault detection and fault type recognition. The organization of this paper is as follows. The primary fault forms of rolling bearings and the major research topics of rolling bearing fault diagnosis are presented in Section 2. Then, Sections 3 and 4 review the typical works on rolling bearing fault detection and fault type recognition, respectively. Section 5 summarizes datasets and the limitations/challenges of existing methods and discusses future research trends in rolling bearing fault diagnosis. Finally, Section 6 draws the conclusions.

## 2. Background, Taxonomy, and Scope

### 2.1. Fault Forms/Types of Rolling Bearing

Rolling bearings have several types, but their basic structures remain the same. Typically, a rolling bearing consists of four parts: the inner ring, the outer ring, the rolling element, and the cage. There are four types of corresponding faults, i.e., the inner ring fault, the outer ring fault, the rolling element fault, and the cage fault. A rolling bearing may fail due to internal and/or external problems/factors. Nowadays, bearing failures are mainly caused by external factors, including improper assembly, oil lubrication failure, pollution corrosion, and overloading. Rolling bearing faults often have the following forms [3]:

(1) Fatigue

Rolling bearings operate with great periodic contact stress between the rolling element and the inner/outer ring surface, causing the contact surface (generally the track surface) to fatigue and crack, which gradually extends to the raceway surface. Fatigue causes the bearing surface material to fall off and form pits. In severe cases, the material on the surface may fall off in large areas. Fatigue pitting and fatigue peeling are commonly used terms for describing fatigue.

(2) Wear

The failures of the rolling bearing sealing system cause bearing wear. When the sealing system fails, foreign matter will enter the bearing, resulting in abnormal friction between the inner ring/outer ring and the rolling elements. Additionally, improper lubrication will further aggravate wear, resulting in continuous material loss, increased surface roughness, increased clearance between bearings, and decreased running accuracy.

(3) Deformation

Deformation means that the bearing surface has undergone plastic deformation, or more specifically, a permanent indentation will appear on the bearing surface if the load borne by the bearing exceeds the yield strength limit of the material. Incorrect assembly methods and foreign matter appearance are the main reasons for the bearing deformation.

(4) Corrosion

Corrosion of rolling bearings occurs when chemical reactions occur on their surface. The first one is the oxidation reaction between the water in the lubricating oil and the bearing surface. The second one is fretting friction between components that leads to the oxidation of surface materials. The last one is abnormal current/voltage that causes local overheating of the bearing, resulting in welding of the element contact surface.

(5) Fracture

Rolling bearing fractures are the damage caused by local stresses exceeding the material's tensile strength limit. Generally, the crack propagates over time and penetrates part of the bearing component, causing complete separation of the material and fracture of the bearing. In addition, violent loading and unloading can also lead to bearing fracture.

### 2.2. Taxonomy and Scope

The purpose of rolling bearing fault diagnosis is to determine the bearing health status by analyzing the collected operation data. Diagnostics of faults revolve primarily around fault detection and fault type identification. Figure 2 shows the general flowchart of fault detection and fault type recognition. Although fault detection and fault type recognition may have some overlap, they are two different types of tasks in fault diagnosis. Specifically, fault detection is to detect faults or non-faults from the collected data, and fault type recognition is to recognize faults and their types from the data. Therefore, to solve these two tasks, different procedures are often used. For fault detection, the collected bearing signals are utilized to determine bearing status. The process often includes removing the noise and harmonic interference from the monitoring signal using signal processing methods and then manually identifying the fault by finding its characteristic frequency. Fault type recognition refers to using the existing bearing signals to construct a diagnostic system to evaluate the unknown bearing signals. Unlike fault detection, fault type recognition methods automatically extract or construct fault features from the signals and determine the bearing health status using machine learning algorithms.

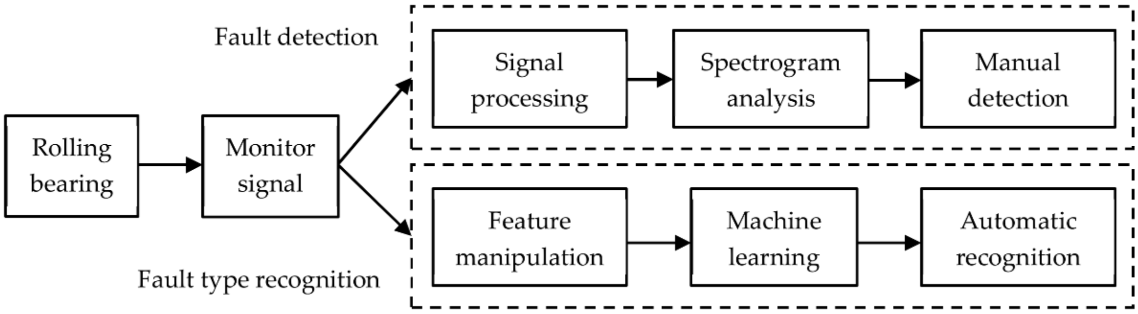



**Figure 2.** General flowchart of rolling bearing fault diagnosis.

The monitoring data of rolling bearings can be collected from oil [11], temperature [12], sound [13], vibration [14], and other media. Performing an oil analysis affects production continuity because it involves shutting down equipment and opening the cover to collect lubricant and other oil samples. Temperature measurement equipment is expensive and cannot provide a promising monitoring effect. Temperature analysis neither has good accuracy at the early stage of bearing fault nor distinguishes the fault types. Sound analysis has high technical demands for signal acquisition and identification because the acoustic signal attenuates and is susceptible to environmental noise interference. In contrast, vibration signal characteristics are stable and easy to collect, making vibration analysis a suitable condition monitoring technique. Vibration analysis has a firm theoretical basis. Research on the fault diagnosis method of rolling bearings based on vibration signal has long been a hot issue concern by domestic and foreign experts and scholars.

This survey paper summarizes the fault diagnosis methods of rolling bearings based on vibration signals from the perspectives of fault detection and fault type recognition. First, four types of signal processing methods commonly used for fault detection of rolling bearings, i.e., morphological transformation-based methods, filter-based methods, decomposition-based methods, and deconvolution-based methods, are discussed. Then, the classical fault type recognition methods are discussed from three aspects: feature extraction, feature reduction, and classification. In addition, the recently popular deep learning based-fault type recognition methods such as convolutional neural networks, Autoencoder, deep belief networks, and recursive neural networks, are also discussed and reviewed. The taxonomy of this survey is shown in Figure 3.

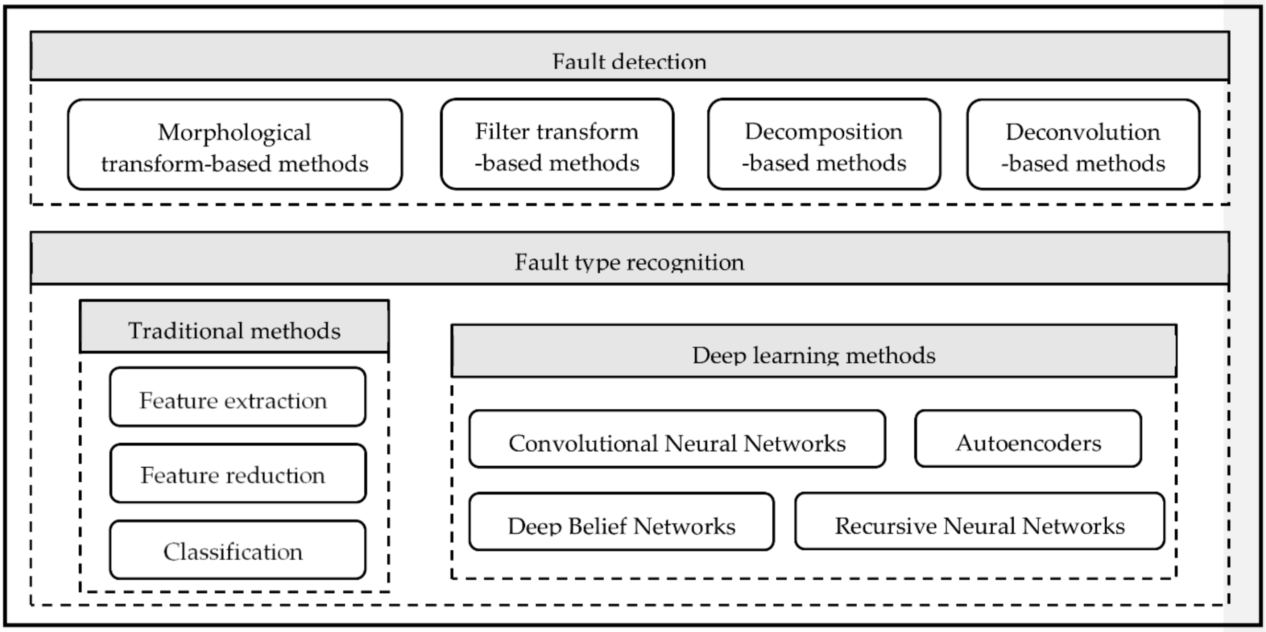



**Figure 3.** The taxonomy of this survey.

## 3. Rolling Bearing Fault Detection

The failure of rolling bearings will break the original energy balance of the system, and the most intuitive performance is abnormal vibration. The bearing fault vibration signal shows an increase or fluctuation in amplitude in the time domain and spectrum lines of fault characteristic frequency with prominent amplitude in the frequency domain. In [15], four empirical formulas were summarized for calculating the theoretical fault frequencies of the inner ring ($F_{inner}$), outer ring ($F_{outer}$), rolling element ($F_{ball}$), and cage ($F_{cage}$), as shown in Equation (1).

$$\begin{cases} F_{inner} = \frac{N_b S_{sh}}{2}(1 - \frac{d_b}{D_p}\cos\varphi) & F_{outer} = \frac{N_b S_{sh}}{2}(1 + \frac{d_b}{D_p}\cos\varphi) \\ F_{ball} = \frac{D_p S_{sh}}{2 d_b}(1 - (\frac{d_b}{D_p}\cos\varphi)^2) & F_{cage} = \frac{S_{sh}}{2}(1 - \frac{d_b}{D_p}\cos\varphi) \end{cases} \quad (1)$$

where $D_p$ is pitch diameter, $d_b$ is rolling element diameter, $N_b$ is rolling element number, $\varphi$ is contact angle, and $S_{sh}$ is shaft speed, which are basic parameters. It is possible to detect a bearing fault by observing the fluctuation of the time-domain waveforms or observing spectral lines associated with the fault characteristic frequency. Directly measuring rolling bearing vibrations is impossible in the real world. Generally, the sensor installed on the bearing pedestal is used to collect the signals indirectly, resulting in a significant amount of noise and harmonic interference in the collected vibration signals. The polluted bearing vibration signal is not effective for detecting bearing faults. Therefore, a series of fault detection methods based on vibration analysis was proposed to remove the noise and harmonic interference components in the signals, enhance the fault-related pulses, reduce the difficulty of fault detection, and improve the effectiveness of detection. Based on the difference in signal processing principles, the fault detection methods are mainly divided into four categories: morphological transformation-based methods, filter-based methods, decomposition-based methods, and deconvolution-based methods.

The common rolling bearing fault detection methods are summarized in Figure 4. The morphological transform-based methods can extract harmonic or impact components of signals by using morphological operators with different structures, whose appropriateness directly influences performance. The filter-based methods can adaptively identify the resonance frequency band that contains rich fault information, where the division of frequency band and the choice of subband are the key factors affecting the results. The decomposition-based methods refer to decomposing the complex signals into simple subband signals, and these methods should address modal aliasing, parameter setting, manually tuning, etc. The deconvolution-based detection methods belong to blind signal processing technology, which recovers fault characteristic signals by designing the appropriate inverse filters and setting the deconvolution period and filter length. In addition to the method based on recursive decomposition, which typically lacks the mathematical model as theoretical support, other methods have the complete mathematical theory.

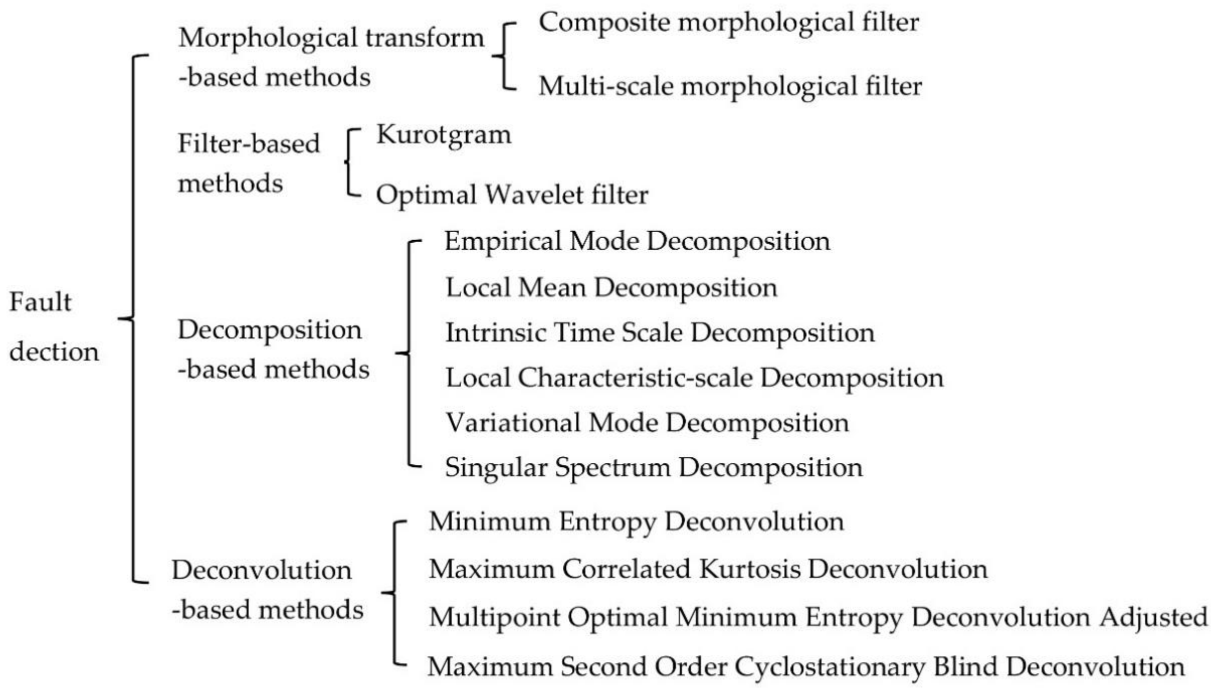



**Figure 4.** Summary of rolling bearing fault detection methods.

### *3.1. Morphological Transform-Based Fault Detection Methods*

The morphological transform-based detection method is a signal processing method based on mathematical morphology theory that can capture the fault-related components in the bearing vibration signals through morphological operators, such as erosion, dilation, open, and close. Matheron introduced mathematical morphology as a denoising method for image processing [16], and then Maragos extended it to the field of signal processing [16,17]. Given the characteristic of morphological transformation to remove signal noise, several researchers have applied it to the fault diagnosis of components of mechanical systems and conducted a great deal of research in recent years. Wang et al. [18] used a morphological close operator to process vibration signals for extracting fault impulses. Shen et al. [19] proposed morphological close–open transform and morphological open–close transform by cascading the close or open operator of the morphology. Li et al. [20] and Raj et al. [21] calculated the gradient (difference) between the dilation and erosion operator of the morphology to obtain the vibration impact component in the signal and defined this procedure as morphological gradient transformation. Following this, some improved methods based on morphological gradient transformation were developed that integrated the close and open operators [22,23] and the close–open and open–close operators [24,25]. These morphological gradient transform methods typically change the negative impact to positive impact, resulting in the change of signal impact components. To ensure that the positive and negative impulses of the signal do not change after morphological transformation, Wang et al. [26] and Meng et al. [27] utilized the mean value operator to fuse the results of the closed and open operators. In addition, Deng et al. [28] and Yan et al. [29] further developed the morphological hat-transform technology, which can enhance the weak impact in the signal by subtracting the morphological transformation result from the original signal.

Recently, Li et al. [30] proposed a morphological gradient product method by multiplying the results of two morphological transforms through a product operator. In addition to developing new morphological transformation methods, multi-scale analysis was used to improve the efficiency of the existing morphological transformation methods [31–38]. To sum up, by analyzing the available morphological transform methods and using cascade operator, gradient operator, hat-transform, product operator, and multi-scale analyses, researchers have developed a series of morphological transform-based fault detection methods with excellent performance.

### 3.2. Filter-Based Fault Detection Methods

The filter-based detection method is to construct a narrow-band filter to remove the noise and interference components from the bearing vibration signals and retain the fault-related pulses. The key to the filter-based method is to determine the center frequency and bandwidth of the narrow-band filter. A typical filter-based detection method is Kurotgram, proposed by Antoni et al. [39,40] in 2006, which uses the bandpass filter of a tree structure to divide the signal spectrum and then calculates the time-domain kurtosis of the filtered signal as a measure of fault information to adaptively select a narrow-band signal with the most fault information for subsequent analysis. The Kurotgram method has two shortcomings: one is that the parameters (center frequency and bandwidth) of the constructed filter are not accurate enough; the other is that the Kurtosis index is easily disturbed by noise, resulting in interference with the selection of the optimal filter. For this reason, Lei et al. [41] and Wang et al. [42] performed a wavelet packet transform on the bearing signal and used each wavelet node as a narrow-band filter to replace the tree structure filter of Kurotgram and proposed two new indicators for evaluating the fault information in the filtered signal, i.e., power spectral kurtosis and power spectral sparsity, to select the optimal filter. Similarly, Chen et al. [43] and Moshrefzadeh et al. [44] used the dual-tree complex wavelet transform and the maximum overlapping discrete wavelet packet transform to generate a series of narrow-band filters, respectively. In addition, many improvements to Kurotgram focus on proposing new evaluation indexes to replace kurtosis, such as spatial spectrum set kurtosis [43], envelope spectrum correlation kurtosis [45], l2/l1 norm [46], negative entropy [47,48], Gini index [49], and weighted cyclic harmonic noise ratio [50], to avoid the wrong selection of filters in the case of excessive non-Gaussian noise or accidental impact.

Although Kurtogram and its improved methods can remove fault-independent noise and harmonic interference from vibration signals, there is still a problem with the accuracy of filter construction, which may lead to the loss of the signal information and affect the extraction of fault-related pulses. As opposed to the traditional method of dividing the frequency band layer by layer, the optimal wavelet filter methods are proposed [51–55]. Tse et al. [51] used the Morlet wavelet as the filter, took maximizing sparsity of the filtered signal as the objective, and applied a genetic algorithm (GA) to locate the center frequency and bandwidth of the optimal Morlet wavelet for automatic filter construction. Similarly, Gu et al. [52] utilized the asymmetric real Laplace wavelet as the filter and determined its center frequency and bandwidth by simultaneously maximizing the impulse and cyclostationary characteristics of the filtered signal.

### 3.3. Decomposition-Based Fault Detection Methods

The decomposition-based detection method involves decomposing the raw vibration signal into several components, such as the fault-related pulse, the noise, and the harmonic interference. Analyzing only the fault-related pulses can simplify the process of detecting the fault. In 1998, Huang et al. [56,57] proposed the empirical mode decomposition (EMD) method, which provides a new idea for analyzing non-stationary signals. Gao [58] utilized EMD to decompose a bearing vibration signal into a series of eigenmode components with inherent oscillation attributes and then conducted envelope spectral analysis to realize bearing fault detection. The EMD method achieved promising performance, but it also has a number of deficiencies, such as end effect, modal aliasing, and over/under envelope, that limit its applications. Huang et al. [59] proposed ensemble empirical mode decomposition (EEMD), where Gaussian white noise is introduced in EMD to assist signal decomposition. Li et al. [60] used EEMD to analyze bearing signals and extract bearing fault features effectively. Even though EEMD can overcome the mode aliasing problem to some degree, it still suffers from problems, such as low decomposition efficiency and the inability to determine white noise amplitude adaptively. Torres et al. [61] proposed complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN), where Gaussian white noise is adaptively added to each stage of the decomposition process. CEEMDAN can improve computing efficiency and reduce construction errors and was successfully applied to detect rolling bearing faults [62–64].

In 2005, Smith et al. [65] proposed the local mean decomposition (LMD), which gradually decomposes a non-stationary signal into a linear combination composed of multiple product function components through the moving average method. In essence, LMD is to separate the pure FM signal and envelope signal from the original signal and multiply the pure FM signal and envelope signal to obtain the product function component with instantaneous frequency and physical significance. LMD shows good performance for bearing fault detection, i.e., can avoid some over/under envelopes and has better signal local characteristics and fewer decomposition components than EMD [65–68]. However, the LMD method can encounter several problems in practical application, such as signal mutation, modal aliasing, and computational inefficiency [69]. In 2007, Frei et al. [70] proposed the intrinsic time scale decomposition (ITD), which can obtain the baseline signal by linear transformation and can adaptively decompose a complex vibration signal into a combination of several proper rotation components (PRCs) and a residual. ITD for bearing fault diagnosis displays significant advantages in end effect, envelope error, and calculation speed over EMD. However, the components decomposed by ITD produce burrs, resulting in distortion of the instantaneous amplitude and frequency [71,72]. Local characteristic-scale decomposition (LCD) was proposed by Cheng et al. [73,74] in 2012, which simultaneously considers the position information of non-stationary signals in the time domain and the frequency domain, avoiding the frequency confusion of EMD and the signal mutation of ITD [75]. Although LCD overcomes the shortcomings of EMD and ITD, there are still some drawbacks, such as the end effect, which often affect the processing results [76].

All the EMD, LMD, ITD, and LCD methods adopt the idea of recursive decomposition, which shares several similar defects. First, the end effect and the mode confusion; Second, the recursive procedure lacks error feedback and correction; Third, the decomposition results are easily affected by noise and abnormal components and have no physical meaning. Dragomiretskiy et al. [77] transformed signal decomposition into a constrained variational problem and proposed variational mode decomposition (VMD), in which the central frequency and bandwidth of each mode depend on the optimal solution variational model found iteratively, avoiding mode aliasing and improving the decomposition accuracy. The decomposition effect of VMD is affected by the number of decomposed modes K and the penalty factor $\alpha$. The particle swarm optimization (PSO) and GA were applied to search the parameter values to enhance the performance of VMD for fault detection [78–80]. Bonizzi et al. [81] proposed the singular spectrum decomposition (SSD), which can adaptively determine the embedding dimension required for each singular value decomposition process and decompose the original signal in narrow-banded components. SSD has the advantages of small end effect, weak mode aliasing and no parameter selection. EMD does not require parameter selection either, but SSD is more effective in decomposing nonlinear and nonstationary time series. There was the development of SSD methods that could improve the decomposition accuracy and the detection ability of weak fault signals, which could be applied more effectively for the fault detection of mechanical equipment [82–84].

### 3.4. Deconvolution-Based Fault Detection Methods

The deconvolution-based detection method is to find an inverse filter to eliminate the transmission path influence in the signal acquisition process and extract the fault pulse from the noise-contaminated vibration signal. The research on deconvolution-based methods can be dated back to 1980. Wiggins et al. [85] proposed the minimum entropy deconvolution (MED) method to maximize the kurtosis of filtered signals and used it to analyze seismic signals. However, MED is easily affected by a random pulse with a large amplitude, making it impossible to accurately extract the periodic pulses corresponding to the fault in the signal [86]. To address this, McDonald et al. [87] developed a new index called correlation kurtosis to evaluate the periodicity and sparsity of signals and proposed maximum correlated kurtosis deconvolution (MCKD) for maximizing the correlation kurtosis value. MCKD overcomes the shortcomings of MED and can effectively extract the periodic pulse corresponding to the fault when there is a single abnormal pulse in the signal. However, the processing performance of MCKD is affected by two parameters, i.e., the inverse filter length and the fault cycle size. Whether the parameter setting is accurate directly affects the final processing result of MCKD. To address this issue, Miao et al. [88] proposed sparse maximum harmonics-noise-ratio deconvolution (SMHD), which can adaptively estimate the fault period by calculating the harmonic noise ratio of the envelope of the filtered signal. However, SMHD generally suffers diminished performance when analyzing the signals with harmonic components. MCKD and SMHD require a long calculation time due to the deconvolution operation based on iteration analysis. Therefore, McDonald et al. [89] proposed a method that does not require iterations, namely, multipoint optimal minimum entropy deconvolution adjusted (MOMEDA), which can complete the deconvolution in a short time but is adversely affected by the periodic oscillations of fault pulse. Recently, Buzzoni et al. [90] introduced the second order cyclostationary index to deconvolution methods and proposed the maximum second order cyclostationary blind deconvolution (CYCBD) method. The performance of CYCBD is better than that of MCKD and MOMEDA, but the fault cycle frequency needs to be set accurately to ensure the processing effect. In order to overcome the shortcomings of these methods, researchers have proposed some improved deconvolution methods by combining other processing methods or using optimization algorithms to determine the optimal parameters required for deconvolution, such as EMD combined with MED [91], PSO optimized MCKD [92], and CS optimized CYCBD [93].

## 4. Rolling Bearing Fault Type Recognition

Unlike the fault detection method, machine learning algorithms were used in the rolling bearing fault type recognition system to replace the manual observation of the fault-related spectral lines. These methods can achieve automatic recognition of different types of faults in rolling bearings.

A summary of the commonly used methods of rolling bearing fault type recognition is shown in Figure 5. These traditional methods need multiple independent steps, such as feature extraction, feature transform or feature selection, and classifier selection and optimization, which often need to be manually set to achieve effective fault recognition performance. The results of the previous step often significantly affect the results of the latter step. Rich domain knowledge is required in the process of fault recognition. The deep-learning-based fault type recognition methods can automatically learn features from the original signals and train classifiers for effective fault recognition without human intervention. However, the deep architecture used in these methods needs rich expertise to design and a large number of samples to train.

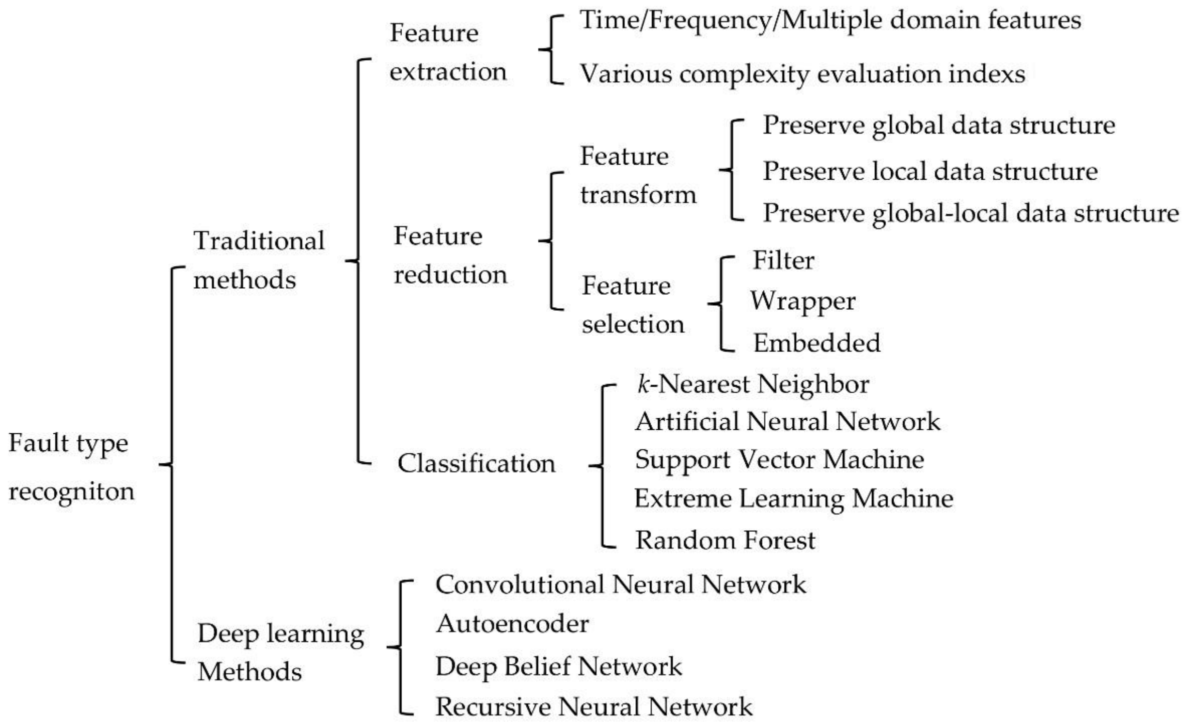



**Figure 5.** Summary of rolling bearing fault type recognition methods.

### *4.1. Traditional Fault Type Recognition Methods*

Traditional rolling bearing fault type recognition methods usually include three steps: feature extraction, feature reduction, and classification.

(1) Feature extraction

Extracting fault-related features from vibration signals is the first step to perform rolling bearing fault type recognition. It is necessary to map the original bearing signals to statistical features to reflect the health status of bearings. Early work on feature extraction of rolling bearing vibration signals mainly focused on calculating various types of time-domain or frequency-domain statistical descriptive indexes [94,95], such as root mean square, kurtosis, skewness, average frequency, and root mean square frequency. These indexes are easy to calculate and intuitive to understand; their values vary with the running state of the rolling bearing.

Complexity can describe the dynamic characteristics of bearing signals under different running conditions. More and more attention is paid to applying various complexity evaluation indexes to fault type recognition. Yang et al. [96] used fractal dimension (FD) to evaluate bearing signals, but the calculation speed of FD is slow, which limits its use in online diagnosis. Caesarendra et al. [97] calculated the Lyapunov exponent of bearing vibration signal as a feature, but its stability is vulnerable to noise interference. The entropy of a time series is an index commonly used to quantify the degree of uncertainty or irregularity. Approximate entropy (AE), sample entropy (SE), fuzzy entropy (FE), permutation entropy (PE), and dispersion entropy (DE) were applied to fault type identification [98–102]. AE has good anti-noise performance when analyzing signals with more data points, but it may cause inaccurate estimation when analyzing signals with fewer data points [103,104]. As an improved form of AE, SE has the advantage of low dependence on the signal length and improved immunity to interference from noise. The disadvantage of SE is that its computation cost is high, and it may not be appropriate for analyzing signals containing similar information [103]. Based on SE, FE introduced a fuzzy membership function and was capable of assessing signal uncertainty more effectively [105]. PE offers simplicity, high robustness to dynamic noise, and a fast calculation speed and can effectively analyze non-stationary signals with complex components [106]. Rostaghi and Azami [107] developed

DE and proved that it was reliable in quantifying the complexity and uncertainty of time signals through the comparative test of various time series. The computational efficiency of DE is significantly better than that of SE, FE, and PE. Instantaneous energy distribution (IED) of bearing vibration signals can describe the time-varying process between fault states. Based on these characteristics, instantaneous energy distribution permutation entropy (IED-PE) [108] and instantaneous energy distribution permutation dispersion entropy (IED-DE) [109] were developed to enhance the accuracy of identifying fault types.

As performing the single-scale analysis of bearing vibration signals with complexity and uncertainty may lead to loss of information, multiscale analysis is introduced to entropy calculation to more accurately evaluate the vibration signals [110–115]. Costa et al. [110] and Aziz et al. [113] used the coarsening method to calculate the entropy of signals on multiple scales and proposed multi-scale sample entropy (MSE) and multi-scale permutation entropy (MPE), respectively. MSE and MPE investigate the irregularity of bearing vibration signals from multiple scales and have made significant progress toward fault diagnosis. In addition, when MSE and MPE are used to analyze signals with fewer data points, their calculation values will fluctuate with the increase in the scale factor, potentially leading to evaluation results instability. Azami et al. [116,117] further proposed multi-scale dispersion entropy (MDE) and refine composite multi-scale dispersion entropy (RCMDE) based on the advantages of DE and the coarse-graining method to address the shortcomings of MSE and MPE. RCMDE and MDE offer much greater computational efficiency than MSE and MPE. Multi-group experiments have shown that RCMDE is more valuable for identifying bearing fault types [117].

(2) Feature reduction

The more statistical features are used to describe signals, the more comprehensively the inherent information of signals is expressed. However, the high-dimensional feature set includes many redundant and negative-effect indexes/features. Dealing with such a large number of useless features typically increases the computational complexity and affects the recognition accuracy. In addition, using too many features to describe a large number of signals may lead to dimensional disaster. Therefore, it is necessary to reduce and compress the tremendous data resources effectively for extracting valuable information and knowledge. Feature transformation and feature selection methods can generate a low-dimensional feature set for fault type recognition.

Feature transformation methods are categorized based on how they preserve data structure. Two types of feature transformation methods exist, the global preservation-based methods, such as principal component analysis (PCA) and linear discriminant analysis (LDA), and the local preservation-based methods, such as local preserving projections (LPP) and margin Fisher analysis (MFA). In [118–121], the sample features reduced by different feature transformation methods were used to perform the identification of bearing fault types. The works in [122–124] show that the feature subset obtained by considering the local and global information in signals with different statuses is more effective for improving recognition performance. Chen et al. [123] proposed the Laplacian LDA (Lap-LDA) method based on least square LDA, which can not only obtain the global structure information of the data using LDA but also obtain the local structure information of the data using the Laplacian map. Zhang et al. [124] proposed global–local structure analysis (GLSA), combining the advantages of LPP and PCA.

Feature selection methods can be divided into three categories, i.e., filter-based methods, wrapper-based methods, and embedded-based methods. In the filter-based method, the features of the original dataset are evaluated and selected according to similarity, dependency, and correlation. This kind of method has fast calculation speed and low complexity. The commonly used methods include Fisher score (FS), Laplacian score (LS), Relief-F, and minimum redundancy maximum relevance (mRMR), which were applied to remove irrelevant features from bearing vibration [125–128]. The wrapper-based feature selection methods use a classifier to evaluate feature subsets for determining the most useful feature subset for classification. Compared with filter-based methods, wrapper-based methods require a longer computing time, but the quality of the feature subset obtained is higher. The wrapper-based feature selection methods could be more efficient through heuristic search algorithms. GA, PSO, and ant colony optimization (ACO) were used for subset search in wrapper-based feature selection methods [129–131]. In order to rapidly and accurately obtain the optimal feature subset for fault type identification, the hybrid feature selection method combining the advantages of the two methods above developed, in which the filter-based method is used as the first selection and the wrapper-based method is used as a second selection [132,133]. The embedded-based methods integrate feature selection and classifier learning, including classification and regression tree (CART) and C4.5 decision tree, which were applied to rolling bearing fault type recognition [134].

(3) Classification

After feature extraction and feature reduction, it is necessary to train a classifier to learn the mapping between the features and the class labels of existing bearing signals for conducting automatic fault type recognition. The known instances with the transformed/selected features and the corresponding class labels are fed into a classification algorithm as the training set. The class label of each instance in the test set can be predicted by the trained classifier according to their features. In the past decade, various classification methods have been applied to rolling bearing fault type recognition, such as $k$-nearest neighbor (KNN), artificial neural network (ANN), support vector machine (SVM), extreme learning machine (ELM), and random forest (RF).

KNN has the advantages of only one parameter and easy implementation by making classification decisions vis identifying the attributes of a limited number of neighboring training samples around the unknown/testing sample. Yan et al. [108] calculated IDE-PE of the bearing signal and used KNN to classify bearing fault types. It should be noted that the performance of KNN depends on the quality of sample features. ANN is a multilayer feedforward neural network and can perform fault type recognition by adjusting the association relationship between a large number of network nodes [135,136]. SVM has good generalization ability, but the kernel function and related parameters need to be selected. Zhu et al. [137] proposed a new rolling element bearing fault diagnosis method based on multi-scale fuzzy entropy, multiple class feature selection, and SVM. Chen et al. [138] input the symbolic entropy of the bearing signals into SVM for fault type identification and obtained good results. ELM is a feedforward neural network that uses random weights between the hidden layer and the input layer, and the output weights of its output layer are calculated through regular processes. With ELM, only the number of hidden layer neurons needs to be set. It has the advantage of rapid processing and good generalization but the disadvantage of overfitting [139]. RF can handle high-dimensional data effectively without a long running time, but the parameter selection of RF often affects the classification accuracy [140]. To avoid setting classifier parameters manually, PSO is used to adaptively determine the optimal parameters of classifiers, e.g., PSO optimized SVM, PSO optimized ELM, and PSO optimized RF, which were proposed to improve the accuracy of fault identification [141–143].

### 4.2. Deep Learning Based Fault Type Recognition Methods

Deep learning techniques have strong learnability. By stacking non-linear processing units layer by layer, it can automatically learn effective features from the raw data without manual feature extraction and manipulation. Deep learning methods are primarily implemented based on ANN, including convolutional neural networks (CNNs), Autoencoder (AE), deep belief networks (DBNs), and recursive neural networks (RNNs). Deep learning methods were used to address machine vision, image processing, speech recognition, text analysis, and other problems. Inspired by these successful applications, deep learning methods have been gradually introduced into the field of fault diagnosis over the past five years [144].

CNNs is the most commonly used deep learning method for fault diagnosis. The network structure is usually composed of convolution layers and pooling layers. The convolution layer convolutes with the original input data to obtain shallow features, and then the pooling layer captures the relatively important features through down sampling. The deep characteristics of the data are gradually obtained by alternately stacking the convolution layer and the pooling layer. CNNs were first used to identify the fault types of rolling bearings in 2016 [145], and then it was widely used and improved [146–148]. The input data of a CNN can be one-dimensional bearing vibration signals or two-dimensional images (i.e., spectrogram, texture, and grayscale) converted from one-dimensional vibration signals. Accordingly, 1D-CNNs and 2D-CNNs methods were developed. Wen et al. [146] transformed the one-dimensional time series signals into two-dimensional image signals through random sampling segments of the original signals and fed these images into Lenet-5, which achieved satisfactory results in three different mechanical fault diagnosis tasks. In [147], Wang et al. applied Morlet wavelet decomposition and bilinear interpolation to convert the vibration signal into grayscale images and then used rectified linear units and the appropriate dropout strategy to improve the generalization performance of CNNs for fault diagnosis. Zhang et al. [148] proposed an improved CNN model using the original vibration signals as inputs. This method uses a wide convolution kernel for extracting features and suppressing high-frequency noise and small convolutional kernels in the preceding layers for performing multilayer nonlinear mapping. The CNN-based fault recognition methods typically extract the internal features of bearing signals through multiple convolution layers and pooling layers and perform fault type recognition by using the fully connected layer, which has a layer with Softmax or Signmoid function for classification, or using other classifiers, such as KNN, to perform classification

AE is a special neural network that consists of two parts, i.e., encoding and decoding, which is to reconstruct input data for obtaining the discriminative data information. The use of improved AE methods has enhanced the processing performance of fault diagnosis. For example, the denoising AE method was proposed by adding noise to the original data, the sparsity AE method was implemented by introducing sparse constraints to the output layer, and the stacking AE method was developed by combining multiple AEs. In [149–152], AE and its improved versions were utilized for extracting discriminative features from the original vibration of signals, based on which bearing fault types may be accurately recognized. Sun et al. [149] used AE to fuse the extracted features of the bearing signals, thereby reducing the redundancy of signals. Shi et al. [150] developed the sparsity AE by adding a sparse penalty to AE for high-level feature learning and bearing fault recognition. Zhou et al. [151] proposed a novel diagnosis method based on Teager computed order spectrum and stacking AE. The results demonstrated that the proposed method could extract features adaptively from bearing vibration signals regardless of the speed or load changes. Gu et al. [152] used a denoising AE to extract features from the bearing original vibration signals and inputted the extracted features to the BP network classifier.

DBN is formed by stacking multiple restricted Boltzmann machines (RBMs), where the output layer of the former RBM is used as the input layer of the latter RBM. These RBMs are trained in a greedy hierarchical manner and can gradually learn expressive features from the data. Oh et al. [153] used the directional gradient histogram of the vibration signals as input features to the DBN model for bearing fault recognition. In [154], the time-domain and frequency-domain features extracted from the different sensor signals were fused as the machine health indicators through a multiple two-layer sparsity AE and used to train a DBN for further classification. Shao et al. [155] developed a novel rolling bearing fault recognition method called continuous DBN with locally linear embedding, which computes a new comprehensive feature index based on locally linear embedding to quantify rolling bearing performance degradation and uses a GA to optimize the DBN parameters for adapting to the signal characteristics.

Considering that the rolling bearing vibration signal is essentially a time series, RNNs with time memory functions have gradually attracted attention. RNNs can effectively analyze and process the time information of the data by establishing the connection between multiple cycle units and mapping the whole history of the input data to the target vector. To address the long-term dependency, improved methods of RNNs were developed, such as long-short-term memory (LSTM) and gated recurrent units (GRUs), which are more effective for bearing fault recognition [156–158]. Yuan et al. [156] investigated the performance of RNN, LSTM, and GRU in fault diagnosis, finding that LSTM performed the best and the ensemble of RNN, LSTM, and GRU could not enhance its performance. Zhao et al. [157] developed convolutional bi-directional LSTM combining CNN and LSTM, where CNN extracted the robust local features from original signals and LSTM encoded temporal information on the outputs of CNN. Zhao et al. [158] constructed a deep GRU for effectively learning features of bearing vibration signals and applied the artificial fish swarm algorithm to obtain the optimal parameters of the GRUs.

## 5. Datasets, Practices, Limitations/Challenges, and Future Research Trends

In this section, commonly used datasets are discussed to provide useful guidance and practices for researchers and practitioners. This section also summarizes the limitations and challenges of existing works and points out future research directions.

### *5.1. Commonly Used Datasets and Practices*

In addition to the development of fault diagnosis methods, the collection and establishment of benchmark datasets are also necessary. The commonly used fault diagnosis datasets are Case Western Reserve University [159], IEEE PHM 2012 Data Challenge [160], University of Cincinnati [161], University of Ottawa [162], and Xi'an Jiao Tong University [163]. They are publicly available and state-of-the-art datasets in the rolling fault diagnosis community. These datasets contain a wide range of rolling bearing operation data, which are described in detail in the corresponding references. For the fault detection problem, a representative signal segment is subjectively intercepted from the collected rolling bearing data for analysis. The detection performance of the same method will vary with different intercepted signals. For the fault type recognition problems, these datasets cannot be directly used to test the effectiveness of the proposed methods. Data preprocessing may be needed to solve the task. For example, the original rolling bearing data of these datasets are often divided to form the training set and the test set to train and test the machine learning-based methods, respectively. In addition to these datasets, there are also some other fault diagnosis datasets that were used in the literature, but they are not publicly available. To make fair comparisons between existing methods, it is important to use the same experimental settings including data preprocessing and splitting. However, this is very hard to achieve at the current stage. On the other hand, to enrich the field of fault diagnosis, it is also necessary to develop/share good datasets of various rolling bearing fault diagnosis tasks, such as the ImageNet [164] dataset in the computer vision community.

### *5.2. Limitations and Challenges*

Although many rolling bearing fault diagnosis methods were proposed and achieved promising results. They have limitations. Most of these methods essentially focus on how to increase the effectiveness of the diagnosis whilst paying little attention to the intelligence and adaptability of the diagnosis systems. Specifically, the limitations/challenges of existing techniques are summarized as follows. Some research directions/topics were also pointed out to address these limitations.

(1) **Limitations of fault detection methods:** Some rolling bearing fault detection methods, such as morphological transform-based methods, filter-based methods, decomposition-based methods, and deconvolution-based methods, often need rich domain/prior knowledge to design and use. For example, it should be known in advance how these methods operate, what their advantages and disadvantages are, and whether they are suitable or effective for the task at hand. However, experts with such knowledge are often costly to employ. In addition, the running condition of rolling bearings in actual services is complex and dynamic, making it very hard to

develop a method to meet the actual environment. Capturing the periodic impact component caused by the fault in the signal is a good way to achieve fault detection but very challenging. To address this limitation, it is promising to develop an intelligent method that can automatically generate a detection model to adaptively remove the background interference and effectively retain the fault-related impulses.

(2) **Limitations of traditional fault type recognition methods:** Traditional rolling bearing fault type recognition methods often include three key steps, i.e., feature extraction, feature reduction, and classification. The results of a previous step may influence the outcomes of the following step. To ensure the whole diagnostic process is feasible and effective, each step must be designed elaborately by experienced researchers, such as determining which type of features to choose/extract, which features to use, which classifier to use, and whether the classifier needs to be optimized. However, it should be noted that such a well-designed diagnostic method may only be effective for a specific fault diagnosis task. Therefore, it is promising to design methods that can automatically deal with these subtasks of fault type recognition. In addition, obtaining representative features of sample signals is the key to achieving good results. Therefore, it is a good research direction that develops a diagnostic method to automatically and simultaneously extract and construct representative features from the original bearing signals, to reduce the difficulty of distinguishing samples and improve the accuracy of fault type recognition.

(3) **Limitations of deep learning-based fault type recognition methods:** Although the deep-learning-based rolling bearing fault type recognition methods can automatically achieve feature extraction, feature reduction, and classification, most of the methods are based on neural networks, which need researchers to design their architectures and adjust the corresponding parameters. The process of model design and parameter adjustment process will consume a significant amount of time and resources. Moreover, the interpretability of the neural network-based methods is not good, i.e., cannot directly express the fault identification process. In addition, these methods usually require a large number of samples to train. However, in practical engineering applications, it is typically difficult to obtain a large number of fault samples, which will limit the use of deep learning-based diagnosis methods.

Therefore, it is necessary to develop new rolling bearing fault type recognition methods that do not need rich manual effort to design the architectures and select the parameters, can effectively deal with limited training data, and learn interpretable models for fault type recognition. These are very challenging research directions, but it is worth investigating them to make the fault type recognition methods more applicable to real-world scenarios.

In summation, the existing rolling bearing fault diagnosis methods require rich prior knowledge and expert experience and lack intelligence and flexibility; therefore, these methods have not been fully explored from a universal perspective. Therefore, it is necessary to develop a rolling bearing fault diagnosis approach that relies less on prior knowledge, domain expert experience, or human intervention and can be effectively applied to a wide range of applications.

### *5.3. Future Research Directions*

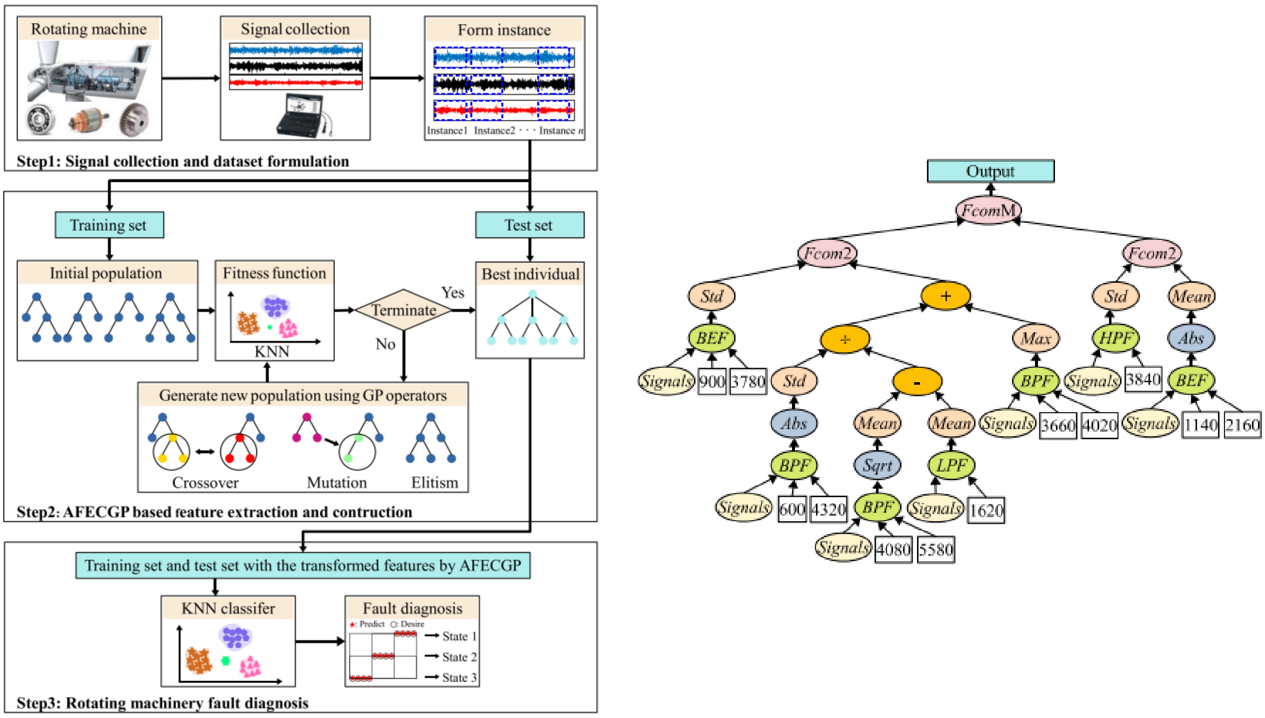
In addition to the aforementioned research directions/topics, there are some other research topics that are becoming popular in this field. This subsection will discuss these research trends.

(1) **Transfer learning-based methods:** The effective performance of the fault type recognition methods usually needs to meet a basic assumption, namely, that the training samples and test samples are independent and identically distributed. However, the monitor information of rolling bearing is generally subject to working conditions, such as the characteristic frequency and amplitude changing with rotational speed, resulting in a large distribution difference between training data and test data, thereby presenting a domain migration issue. Transfer learning (TL) can extract knowledge

from one or more related scenes to help improve the learning performance of scenarios in the target domain [165]. TL can relax the assumption of independent and identical distributions and provide a new solution to address the above deficiencies. The TL-based rolling bearing fault type recognition methods were proposed and achieved desirable results [166–168]. The TL-based recognition model, learning the common feature space from the source domain data and the target domain data to reduce the distribution difference between different domains, cannot adaptively adjust its parameters for target domain tasks, thereby affecting its domain adaptability and recognition accuracy. Thus, the further development of TL-based fault type recognition methods is a good direction for future research to improve the classification performance, recognition accuracy, and generalization under variable operating conditions.

(2) **Few-shot learning methods:** A large amount of labelled data is also the key to ensuring the performance of existing fault type recognition methods, especially for deep learning-based methods. In real-world scenarios, it is easy to obtain enough normal samples due to the rolling bearing mostly running under normal conditions, but the fault samples are typically difficult to obtain and require extensive manual effort to label. The absence of labelled fault samples will either lead to overfitting in the training process or the class imbalance problem. Few-shot learning (FSL) is effective for distinguishing failure attribution accurately under very limited data conditions [169,170]. Data augmentation, data/model transfer, and meta-learning constitute the three main threads of FSL methods. Thus, the comprehensive exploration of FSL-based fault type recognition methods is a good direction for future research for reducing the dependence on large amounts of data, avoiding the risk of overfitting, and improving the applicability and recognition performance.

(3) **Evolutionary deep learning methods:** Evolutionary deep learning methods aim to deal with the limitations of deep learning methods, particularly neural networks, by using evolutionary computation (EC) techniques. This direction includes two main topics, i.e., using EC methods to automatically design neural networks and using EC methods to evolve deep models by themselves. On the first topic, some work was performed to evolve neural networks for fault diagnosis by finding the optimal numbers of layers, network connections, numbers of filters, etc. [171–175]. These methods can reduce the requirement of expertise from both the neural network domain and the problem domain, improve recognition performance, and decrease the number of parameters in the evolved models. On the second topic, pure EC methods, particularly genetic programming methods, are used to evolve deep models. GP is a computational intelligence algorithm to achieve automatic programming without human intervention and domain knowledge [176,177]. With a flexible program expression, GP can automatically evolve variable-length models to solve a task. GP has shown promise in the computer vision domain by evolving deep models [178–181]. The models evolved by GP typically have better interpretability than neural networks. However, there is little work on GP for fault diagnosis [182–184]. Figure 6 shows an example of using GP to solve fault type recognition, where the GP method is used to automatically generate informative and discriminative features from original vibration signals for recognizing different fault types. The left example tree of Figure 6 is the solution evolved by GP, showing high interpretability. In addition, the solutions are often creative and even not considered by human experts [183,184]. However, both topics have not been fully investigated in the fault diagnosis community. Therefore, it is promising to develop effective evolutionary deep learning approaches to fault diagnosis.




(**a**) Flowchart (**b**) Solution

**Figure 6.** Illustrations of the proposed GPAFEC method in [183]. (**a**) Flowchart of GPAFEC, (**b**) Solution evolved by GP.

## 6. Conclusions

The rolling bearing is an indispensable part of rotating machinery, and its running status typically affects the operation of the whole equipment. The research into rolling bearing fault diagnosis technology is of great significance to ensure the safe and stable operations of rotating machinery. This paper comprehensively reviewed existing fault diagnosis methods of the rolling bearing in terms of fault detection and fault type recognition. For fault detection, the methods, i.e., morphological transformation-based methods, filter-based methods, decomposition-based methods, and deconvolution-based methods, were discussed. For fault type recognition, traditional methods and deep learning-based methods were discussed. The commonly used datasets of fault diagnosis were presented for better practices. In addition, we summarized the limitations of existing methods and pointed out future research directions, which provides helpful guidance for researchers who are interested in this field. Overall, this field of fault diagnosis has potential for future study. Given the current limitations, it is still needed to develop automatic, intelligent, effective, and efficient methods for rolling bearing fault diagnosis under real-world scenarios. In addition, some topics such as transfer learning, few-shot learning, and evolutionary deep learning can also be further investigated to enrich this field.

**Author Contributions:** B.P.: writing—original draft preparation; Y.B.: writing—review and editing; B.X., M.Z. and S.W.: supervision and revision. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Scientific Research Foundation for the Talents of Hebei Agricultural University (No: YJ2021056) and Hebei Key Laboratory of Electric Machinery Health Maintenance and Failure Prevention Fund (No.: KF2021-01).

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References


1. Chen, S.; Peng, Z.; Zhou, P. Review of signal decomposition theory and its applications in machine fault diagnosis. *J. Mech. Eng.* **2020**, *56*, 91–107.
2. Yan, G.; Chen, J.; Bai, Y.; Yu, C.; Yu, C. A Survey on Fault Diagnosis Approaches for Rolling Bearings of Railway Vehicles. *Processes* **2022**, *10*, 724. [CrossRef]
3. Kuang, P.; Xu, F.; Liu, Y. *Modern Machinery Fault Diagnosis: Principles and Techniques*; China Agriculture Press: Beijing, China, 1991.
4. Wang, X. Research on Fault Diagnosis Method of Rolling Bearing Based on Vibration Signal Processing. Ph.D. Thesis, North China Electric Power University, Beijing, China, 2017.
5. Zhang, X.; Zhao, B.; Lin, Y. Machine Learning Based Bearing Fault Diagnosis Using the Case Western Reserve University Data: A Review. *IEEE Access* **2021**, *9*, 155598–155608. [CrossRef]
6. Singh, J.; Azamfar, M.; Li, F.; Lee, J. A systematic review of machine learning algorithms for prognostics and health management of rolling element bearings: Fundamentals, concepts and applications. *Meas. Sci. Technol.* **2020**, *32*, 012001. [CrossRef]
7. Lin, H.; Ye, Y. Reviews of bearing vibration measurement using fast Fourier transform and enhanced fast Fourier transform algorithms. *Adv. Mech. Eng.* **2019**, *11*, 1687814018816751. [CrossRef]
8. Liu, R.; Yang, B.; Zio, E.; Chen, X. Artificial intelligence for fault diagnosis of rotating machinery: A review. *Mech. Syst. Signal Process.* **2018**, *108*, 33–47. [CrossRef]
9. Wang, Y.; Xiang, J.; Markert, R.; Liang, M. Spectral kurtosis for fault detection, diagnosis and prognostics of rotating machines: A review with applications. *Mech. Syst. Signal Process.* **2016**, *66*, 679–698. [CrossRef]
10. Rai, A.; Upadhyay, S. A review on signal processing techniques utilized in the fault diagnosis of rolling element bearings. *Tribol. Int.* **2016**, *96*, 289–306. [CrossRef]
11. Neupane, D.; Seok, J. Bearing fault detection and diagnosis using case western reserve university dataset with deep learning approaches: A review. *IEEE Access* **2020**, *8*, 93155–93178. [CrossRef]
12. Yang, Z. Oil Liquid State Monitoring technolgoy and application in equipment maintenance management. *Mech. Manag. Dev.* **2006**, *89*, 51–52.
13. Sun, B.; Wang, Y.; Yang, L. Study of fault diagnosis of induction motor bearing based on infrared inspection. *Elect. Mach. Control* **2012**, *16*, 50–55.
14. Wang, Y. Acoustic-Based Condition Monitoring of Machinery Using Blind Signal Processing. Ph.D. Thesis, Kunming University of Science and Technology, Kunming, China, 2010.
15. Singh, S.; Vishwakarma, M. A review of vibration analysis techniques for rotating machines. *Int. J. Eng. Res. T.* **2015**, *4*, 757–761.
16. Randall, R.; Jérme, A. Rolling element bearing diagnostics—A tutorial. *Mech. Syst. Signal Process.* **2011**, *25*, 485–520. [CrossRef]
17. Ripley, B.; Matheron, G. Random sets and integral geometry. *J. R. Stat. Soc.* **1975**, *139*, 277–278. [CrossRef]
18. Maragos, P.; Schafer, R. Morphological filters. Part 1. Their set-theoretic analysis and relations to linear shift-invariant filters. *IEEE Trans. Acous. Speech Signal Process.* **1987**, *35*, 1153–1169. [CrossRef]
19. Maragos, P.; Schafer, R. Morphological filters. Part 2. Their set-theoretic analysis and relations to linear shift-invariant filters. *IEEE Trans. Acous. Speech Signal Process.* **1987**, *35*, 1170–1184. [CrossRef]
20. Wang, J.; Xu, G.; Zhang, Q.; Liang, L. Application of improved morphological filter to the extraction of impulsive attenuation signals. *Mech. Syst. Signal Process.* **2009**, *23*, 236–245. [CrossRef]
21. Shen, C.; Zhu, Z.; Kong, F.; Huang, W. An improved morphological filtering method and its application in bearing fault feature extraction. *J. Vib. Eng.* **2012**, *25*, 468–473.
22. He, W.; Jiang, Z.; Qin, Q. A joint adaptive wavelet filter and morphological signal processing method for weak mechanical impulse extraction. *J. Mech. Sci. Technol.* **2010**, *24*, 1709–1716. [CrossRef]
23. Raj, A.; Murali, N. Early classification of bearing faults using morphological operators and fuzzy inference. *IEEE Trans. Ind. Electron.* **2013**, *60*, 567–574. [CrossRef]
24. Osman, S.; Wang, W. An Hilbert-huang spectrum technique for fault detection in rolling element bearings. *IEEE Trans. Instrum. Meas.* **2016**, *65*, 2646–2656. [CrossRef]
25. Li, Y.; Liang, X.; Zuo, M. A new strategy of using a time-varying structure element for mathematical morphological filtering. *Measurement* **2017**, *106*, 53–65. [CrossRef]
26. Li, Y.; Zuo, M.; Lin, J.; Liu, J. Fault detection method for railway wheel flat using an adaptive multiscale morphological filter. *Mech. Syst. Signal Process.* **2017**, *84*, 642–658. [CrossRef]
27. Li, Y.; Liang, X.; Zuo, M. Diagonal slice spectrum assisted optimal scale morphological filter for rolling element bearing fault diagnosis. *Mech. Syst. Signal Process.* **2017**, *85*, 146–161. [CrossRef]
28. Wang, D.; Tse, P.; Tse, Y. A morphogram with the optimal selection of parameters used in morphological analysis for enhancing the ability in bearing fault diagnosis. *Meas. Sci. Technol.* **2012**, *23*, 65001–65015. [CrossRef]
29. Meng, L.; Xiang, J.; Wang, Y.; Jiang, Y.; Gao, H. A hybrid fault diagnosis method using morphological filter-translation invariant wavelet and improved ensemble empirical mode decomposition. *Mech. Syst. Signal Process.* **2015**, *50–51*, 101–115. [CrossRef]
30. Deng, F.; Tang, G.; He, Y. Fault feature extraction for rolling element bearings based on cepstrum pre-whitening and morphology self-complementary top-hat transformation. *J. Vib. Shock* **2015**, *34*, 77–81.
31. Yan, X.; Jia, M. Parameter optimized combination morphological filter-hat transform and its application in fault diagnosis of wind turbine. *J. Mech. Eng.* **2016**, *52*, 103–110. [CrossRef]

32. Li, Y.; Zuo, M.; Chen, Y.; Feng, K. An enhanced morphology gradient product filter for bearing fault detection. *Mech. Syst. Signal Process.* **2018**, *109*, 166–184. [CrossRef]
33. Deng, F.; Yang, S.; Guo, W.; Liu, Y. Fault feature extraction method for rolling bearing based on adaptive multi-scale morphological AVG-Hat filtering. *J. Vib. Eng.* **2017**, *30*, 178–187.
34. Zou, F.; Zhang, H.; Sang, S.; Li, X.; He, W.; Liu, X. Bearing fault diagnosis based on combined multi-scale weighted entropy morphological filtering and bi-LSTM. *Appl. Intell.* **2021**, *51*, 6647–6664. [CrossRef]
35. Li, Y.; Liang, X.; Lin, J.; Chen, Y.; Liu, J. Train axle bearing fault detection using a feature selection scheme based multi-scale morphological filter. *Mech. Syst. Signal Process.* **2018**, *101*, 435–448. [CrossRef]
36. Zhu, D.; Zhang, Y.; Zhu, Q. Fault feature extraction for rolling element bearings based on multi-scale morphological filter and frequency-weighted energy operator. *J. Vibroeng.* **2018**, *20*, 2892–2907. [CrossRef]
37. Wu, Z.; Yang, S.; Ren, B.; Ma, X.; Zhang, J. Rolling element bearing fault diagnosis method based on NAMEMD and multi-scale morphology. *J. Vib. Shock* **2016**, *35*, 127–133.
38. Chen, Q.; Chen, Z.; Sun, W.; Yang, G. A new structuring element for multi-scale morphology analysis and its application in rolling element bearing fault diagnosis. *J. Vib. Control* **2015**, *21*, 765–789. [CrossRef]
39. Antoni, J. The spectral kurtosis: A useful tool for characterising non-stationary signals. *Mech. Syst. Signal Process.* **2006**, *20*, 282–307. [CrossRef]
40. Antoni, J.; Randall, R. The spectral kurtosis: Application to the vibratory surveillance and diagnostics of rotating machines. *Mech. Syst. Signal Process.* **2006**, *20*, 308–331. [CrossRef]
41. Lei, Y.; Lin, J.; He, Z.; Zi, Y. Application of an improved kurtogram method for fault diagnosis of rolling element bearings. *Mech. Syst. Signal Process.* **2011**, *25*, 1738–1749. [CrossRef]
42. Wang, D.; Tse, P.; Tsui, K. An enhanced kurtogram method for fault diagnosis of rolling element bearings. *Mech. Syst. Signal Process.* **2013**, *35*, 176–199. [CrossRef]
43. Chen, B.; Zhang, Z.; Zi, Y.; He, Z.; Sun, C. Detecting of transient vibration signatures using an improved fast spatial-spectral ensemble kurtosis kurtogram and its applications to mechanical signature analysis of short duration data from rotating machinery. *Mech. Syst. Signal Process.* **2013**, *40*, 1–37. [CrossRef]
44. Moshrefzadeh, A.; Fasana, A. The autogram: An effective approach for selecting the optimal demodulation band in rolling element bearings diagnosis. *Mech. Syst. Signal Process.* **2018**, *105*, 294–318. [CrossRef]
45. Gu, X.; Yang, S.; Liu, Y.; Liao, Y. An improved kurtogram method and its application in fault diagnosis of rolling element bearings under complex interferences. *J. Vib. Shock* **2017**, *36*, 187–193.
46. Tse, P.; Wang, D. The design of a new sparsogram for fast bearing fault diagnosis: Part 1 of the two related manuscripts that have a joint title as "Two automatic vibration-based fault diagnostic methods using the novel sparsity measurement-Parts 1 and 2". *Mech. Syst. Signal Process.* **2013**, *40*, 499–519. [CrossRef]
47. Antoni, J. The infogram: Entropic evidence of the signature of repetitive transients. *Mech. Syst. Signal Process.* **2016**, *74*, 73–94. [CrossRef]
48. Wan, S.; Zhang, X.; Dou, L. Shannon entropy of binary wavelet packet subbands and its application in bearing fault extraction. *Entropy* **2018**, *20*, 260. [CrossRef]
49. Miao, Y.; Zhao, M.; Lin, J. Improvement of kurtosis-guided-grams via Gini index for bearing fault feature identification. *Meas. Sci. Technol.* **2017**, *28*, 125001. [CrossRef]
50. Mo, Z.; Wang, J.; Zhang, H.; Miao, Q. Weighted cyclic harmonic-to-noise ratio for rolling element bearing fault diagnosis. *IEEE Trans. Instrum. Meas.* **2020**, *69*, 432–442. [CrossRef]
51. Tse, P.; Wang, D. The automatic selection of an optimal wavelet filter and its enhancement by the new sparsogram for bearing fault detection. *Mech. Syst. Signal Process.* **2013**, *40*, 520–544. [CrossRef]
52. Gu, X.; Yang, S.; Liu, Y.; Ren, B.; Zhang, J. Fault Feature Extraction of Wheel-bearing Based on Multi-objective Cross Entropy Optimization. *J. Mech. Eng.* **2018**, *54*, 304–311. [CrossRef]
53. Wan, S.; Peng, B. Adaptive asymmetric real Laplace wavelet filtering and its application on rolling bearing early fault diagnosis. *Shock Vib.* **2019**, *2019*, 7475868. [CrossRef]
54. Xu, Y.; Zhang, K.; Ma, C.; Sheng, Z.; Shen, H. An adaptive spectrum segmentation method to optimize empirical wavelet transform for rolling bearings fault diagnosis. *IEEE Access* **2019**, *7*, 30437–30456. [CrossRef]
55. Guo, J.; Shi, Z.; Zhen, D.; Meng, Z.; Gu, F.; Ball, A.D. Modulation signal bispectrum with optimized wavelet packet denoising for rolling bearing fault diagnosis. *Struct. Health Monit.* **2022**, *21*, 984–1011. [CrossRef]
56. Huang, N.; Shen, Z.; Long, S.; Wu, M.; Shih, H.; Zheng, Q.; Yen, N.; Tung, C.; Liu, H. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. R. Soc. Lond. Ser. Math. Phys. Eng. Sci.* **1998**, *454*, 903–995. [CrossRef]
57. Huang, N.; Zheng, S.; Long, S. A new view of nonlinear water waves: The Hilbert spectrum. *Annu. Rev. Fluid Mech.* **1999**, *31*, 417–457. [CrossRef]
58. Gao, Q.; Du, X.; Fan, H.; Meng, Q. An empirical mode decomposition based method for rolling bearing fault diagnosis. *J. Vib. Eng.* **2007**, *20*, 19–22.
59. Wu, Z.; Huang, N. Ensemble empirical mode decomposition: A noise-assisted data analysis method. *Adv. Adapt. Data Anal.* **2009**, *1*, 1–41. [CrossRef]

60. Li, H.; Liu, T.; Wu, X.; Chen, Q. Application of EEMD and improved frequency band entropy in bearing fault feature extraction. *ISA Trans.* **2019**, *88*, 170–185. [CrossRef]
61. Tomes, M.; Colominas, M.; Schlotthauer, G.; Flandrin, P. A complete ensemble empirical mode decomposition with adaptive noise. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, Prague, Czech Republic, 22–27 May 2011.
62. Rabah, A.; Abdelhafid, K.; Azeddine, B.; Derouiche, Z. Rolling bearing fault diagnosis based on an improved denoising method using the complete ensemble empirical mode decomposition and the optimized thresholding operation. *IEEE Sens. J.* **2018**, *18*, 7166–7172.
63. Huang, H.; Sun, S.; Ren, X.; Liu, H. Early fault diagnosis of rolling bearing based on CEEMDAN and 1.5 dimension spectrum. *China Meas. Test* **2019**, *4*, 155–160.
64. Gao, S.; Wang, Q.; Zhang, Y. Rolling bearing fault diagnosis based on CEEMDAN and refined composite multiscale fuzzy entropy. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 1–8. [CrossRef]
65. Smith, S. The local mean decomposition and its application to EEG perception data. *J. R. Soc. Interface* **2005**, *2*, 443–454. [CrossRef]
66. Cheng, J.; Yang, Y.; Yang, Y. A rotating machinery fault diagnosis method based on local mean decomposition. *Digital Signal Process.* **2012**, *22*, 356–366. [CrossRef]
67. Wang, L.; Liu, Z.; Miao, Q.; Zhang, X. Complete ensemble local mean decomposition with adaptive noise and its application to fault diagnosis for rolling bearings. *Mech. Syst. Signal Process.* **2018**, *106*, 24–39. [CrossRef]
68. Xu, Y.; Zhang, K.; Ma, C.; Li, S.; Zhang, H. Optimized LMD method and its applications in rolling bearing fault diagnosis. *Meas. Sci. Technol.* **2019**, *30*, 125017. [CrossRef]
69. Li, X.; Ma, J.; Wang, X.; Wu, J.; Li, Z. An improved local mean decomposition method based on improved composite interpolation envelope and its application in bearing fault feature extraction. *ISA Trans.* **2020**, *97*, 365–383. [CrossRef]
70. Frei, M.; Osorio, I. Intrinsic time-scale decomposition: Time-frequency-energy analysis and real-time filtering of non-stationary signals. *Proc. Math. Phys. Eng. Sci.* **2007**, *463*, 321–342. [CrossRef]
71. Yu, J.; Liu, H. Sparse coding shrinkage in intrinsic time-scale decomposition for weak fault feature extraction of bearings. *IEEE Trans. Instrum. Meas.* **2018**, *67*, 1579–1592. [CrossRef]
72. Ma, J.; Zhan, L.; Li, C.; Li, Z. An improved intrinsic time-scale decomposition method based on adaptive noise and its application in bearing fault feature extraction. *Meas. Sci. Technol.* **2020**, *32*, 025103. [CrossRef]
73. Yang, Y.; Zeng, M.; Cheng, J. A New Time-frequency analysis method-the local characteristic-scale decomposition. *J. Hunan Univ. (Nat. Sci.)* **2012**, *39*, 35–39.
74. Cheng, J.; Yang, Y.; Yang, Y. Local characteristic-scale decomposition method and its application to gear fault diagnosis. *J. Mech. Eng.* **2012**, *48*, 64–71. [CrossRef]
75. Cheng, J.; Yang, Y.; Li, X.; Pan, H.; Cheng, J. An early fault diagnosis method of gear based on improved symplectic geometry mode decomposition. *Measurement* **2019**, *151*, 107140. [CrossRef]
76. Luo, S.; Yang, W.; Luo, Y. A novel fault detection scheme using improved inherent multiscale fuzzy entropy with partly ensemble local characteristic-scale decomposition. *IEEE Access* **2020**, *8*, 6650–6661. [CrossRef]
77. Dragomiretskiy, K.; Zosso, D. Variational mode decomposition. *IEEE Trans. Signal Process.* **2014**, *62*, 531–544. [CrossRef]
78. Bian, J. Fault Diagnosis of bearing combining parameter optimized variational mode decomposition based on genetic algorithm with 1.5-dimensional spectrum. *J. Propul. Technol.* **2017**, *38*, 1618–1624.
79. Yan, X.; Jia, M. Application of CSA-VMD and optimal scale morphological slice bispectrum in enhancing outer race fault detection of rolling element bearings. *Mech. Syst. Signal Process.* **2019**, *122*, 56–86. [CrossRef]
80. Li, H.; Liu, T.; Wu, X.; Chen, Q. An optimized VMD method and its applications in bearing fault diagnosis. *Measurement* **2020**, *166*, 108185. [CrossRef]
81. Bonizzi, P.; Karel, J.; Meste, O.; Peeters, R. Singular spectrum decomposition: A new method for time series decomposition. *Adv. Adapt. Data Anal.* **2014**, *6*, 107–109. [CrossRef]
82. Xu, Y.; Zhang, Z.; Ma, C.; Zhang, J. Improved singular spectrum decomposition and its applications in rolling bearing fault diagnosis. *J. Vib. Eng.* **2019**, *32*, 168–175.
83. Wang, X.; Tang, G.; He, Y. Weak fault diagnosis for rolling bearing based on COT-SSD under variable rotating speed. *Elec. Power Autom. Equip.* **2019**, *39*, 187–193.
84. Mao, Y.; Jia, M.; Yan, X. A new bearing weak fault diagnosis method based on improved singular spectrum decomposition and frequency-weighted energy slice bispectrum. *Measurement* **2020**, *166*, 108235. [CrossRef]
85. Wiggins, R. Minimum entropy deconvolution. *Geophys. Prospect. Petrole* **1980**, *16*, 21–35. [CrossRef]
86. Endo, H.; Randall, R. Enhancement of autoregressive model based gear tooth fault detection technique by the use of minimum entropy deconvolution filter. *Mech. Syst. Signal Process.* **2007**, *21*, 906–919. [CrossRef]
87. McDonald, G.; Zhao, Q.; Zuo, M. Maximum correlated kurtosis deconvolution and application on gear tooth chip fault detection. *Mech. Syst. Signal Process.* **2012**, *33*, 237–255. [CrossRef]
88. Miao, Y.; Zhao, M.; Lin, J.; Xu, X. Sparse maximum harmonics-to-noise-ratio deconvolution for weak fault signature detection in bearings. *Meas. Sci. Technol.* **2016**, *27*, 105004. [CrossRef]
89. McDonald, G.; Zhao, Q. Multipoint optimal minimum entropy deconvolution and convolution fix: Application to vibration fault detection. *Mech. Syst. Signal Process.* **2017**, *82*, 461–477. [CrossRef]

90. Buzzonia, M.; Antoni, J.; D'Elia, G. Blind deconvolution based on cyclostationarity maximization and its application to fault identification. *J. Sound Vib.* **2018**, *432*, 569–601. [CrossRef]
91. Zhang, Z.; Entezami, M.; Stewart, E.; Roberts, C. Enhanced fault diagnosis of roller bearing elements using a combination of empirical mode decomposition and minimum entropy deconvolution. *Proc. Inst. Mech. Eng. Part C J. Mech. Eng. Sci.* **2017**, *231*, 655–671. [CrossRef]
92. Cheng, Y.; Wang, Z.; Zhang, W.; Huang, G. Particle swarm optimization algorithm to solve the deconvolution problem for rolling element bearing fault diagnosis. *ISA Trans.* **2019**, *90*, 244–267. [CrossRef]
93. Wang, X.; Yan, X.; He, Y. Weak fault detection for wind turbine bearing based on ACYCBD and IESB. *J. Mech. Sci. Technol.* **2020**, *34*, 1399–1413. [CrossRef]
94. Hu, Q.; He, Z.; Zhang, S.; Zi, Y.; Lei, Y. Intelligent diagnosis for incipient fault based on lifting wavelet package transform and support vector machines ensemble. *J. Mech. Eng.* **2006**, *42*, 20–26. [CrossRef]
95. Lei, Y.; He, Z.; Zi, Y. Fault diagnosis based on novel hybrid intelligent model. *J. Mech. Eng.* **2008**, *44*, 112–117. [CrossRef]
96. Yang, J.; Zhang, Y.; Zhu, Y. Intelligent fault diagnosis of rolling element bearing based on SVMs and fractal dimension. *Mech. Syst. Signal Process.* **2007**, *21*, 2012–2024. [CrossRef]
97. Caesarendra, W.; Kosasih, B.; Tieu, A.; Moodie, C. Application of the largest Lyapunov exponent algorithm for feature extraction in low speed slew bearing condition monitoring. *Mech. Syst. Signal Process.* **2015**, *50–51*, 116–138. [CrossRef]
98. Yan, R.; Gao, R. Approximate entropy as a diagnostic tool for machine health monitoring. *Mech. Syst. Signal Process.* **2007**, *21*, 824–839. [CrossRef]
99. Su, W.; Wang, F.; Zhu, H.; Guo, Z.; Zhang, Z.; Zhang, H. Feature extraction of rolling element bearing fault using wavelet packet sample entropy. *J. Vib. Meas. Diag.* **2011**, *31*, 33–37+134.
100. Zheng, J.; Cheng, J.; Yang, Y. A rolling bearing fault diagnosis approach based on LCD and fuzzy entropy. *Mech. Mach. Theory* **2013**, *70*, 441–453. [CrossRef]
101. Yan, R.; Liu, Y.; Gao, R. Permutation entropy: A nonlinear statistical measure for status characterization of rotary machines. *Mech. Syst. Signal Process.* **2012**, *29*, 474–484. [CrossRef]
102. Fu, W.; Tang, J.; Wang, K. Semi-supervised fault diagnosis of bearings based on the VMD dispersion entropy and improved SVDD with modified grey wolf optimizer. *J. Vib. Shock* **2019**, *38*, 190–197.
103. Pincus, S. Approximate entropy (ApEn) as a complexity measure. *Chaos* **1998**, *5*, 110–117. [CrossRef]
104. Richman, J.; Randall, M. Physiological time-series analysis using approximate entropy and sample entropy. *Am. J. Physiol. Heart C.* **2000**, *278*, 2039–2049. [CrossRef]
105. Chen, W. A Study of Feature Extraction from sEMG Singal Based on Entropy. Ph.D. Thesis, Shanghai University, Shanghai, China, 2008.
106. Bandt, C.; Pompe, B. Permutation entropy: A natural complexity measure for time series. *Phys. Rev. Lett.* **2002**, *88*, 174102. [CrossRef] [PubMed]
107. Rostaghi, M.; Azami, H. Dispersion entropy: A measure for time-series analysis. *IEEE Signal Process. Lett.* **2016**, *23*, 610–614. [CrossRef]
108. Yan, X.; Jia, M.; Zhao, Z. A novel intelligent detection method for rolling bearing based on IVMD and instantaneous energy distribution-permutation entropy. *Measurement* **2018**, *130*, 435–447. [CrossRef]
109. Tang, G.; Pang, B.; He, Y.; Tian, T. Gearbox fault diagnosis based on hierarchical instantaneous energy density dispersion entropy and dynamic time warping. *Entropy* **2019**, *21*, 593. [CrossRef]
110. Costa, M.; Goldberger, A.; Peng, C. Multiscale entropy analysis of complex physiologic time series. *Phys. Rev. Lett.* **2007**, *89*, 705–708. [CrossRef]
111. Liu, H.; Han, M. A fault diagnosis method based on local mean decomposition and multi-scale entropy for roller bearings. *Mech. Mach. Theory* **2014**, *75*, 67–78. [CrossRef]
112. Zhang, L.; Huang, W.; Xiong, G. Assessment of rolling element bearing fault severity using multi-scale entropy. *J. Vib. Shock* **2014**, *33*, 185–189.
113. Aziz, W.; Arif, M. Multiscale permutation entropy of physiological time series. In Proceedings of the INMIC 2005 9th International Multitopic Conference, Karachi, Pakistan, 1–5 December 2005.
114. Tiwari, R.; Gupta, V.; Kankar, P. Bearing fault diagnosis based on multi-scale permutation entropy and adaptive neuro fuzzy classifier. *J. Vib. Control* **2015**, *21*, 461–467. [CrossRef]
115. Zheng, J.; Cheng, J.; Yang, Y. Multi-scale Permutation entropy and its applications to rolling bearing fault diagnosis. *China Mech. Eng.* **2013**, *24*, 2641–2646.
116. Azami, H.; Kinney-Lang, E.; Ebied, A.; Fernández, A.; Escudero, J. Multiscale dispersion entropy for the regional analysis of resting-state magnetoencephalogram complexity in Alzheimer's disease. In Proceedings of the 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Jeju Island, South Korea, 11–15 July 2017.
117. Azami, H.; Rostaghi, M.; Abasolo, D.; Escudero, J. Refined composite multiscale dispersion entropy and its application to biomedical signals. *IEEE Trans. Biomed. Eng.* **2017**, *64*, 2872–2879.
118. Xu, Z.; Liu, K.; Zhang, H.; Wnag, D.; Zhang, M. A fault diagnosis method for rolling bearings based on empirical mode decomposition and principal component analysis. *J. Vib. Shock* **2014**, *33*, 133–139.

119. Ahmed, H.; Nandi, A. Three-stage hybrid fault diagnosis for rolling bearings with compressively-sampled data and subspace learning techniques. *IEEE Trans. Ind. Electron.* **2018**, *66*, 5516–5524. [CrossRef]
120. Ding, X.; He, Q.; Luo, N. A fusion feature and its improvement based on locality preserving projections for rolling element bearing fault classification. *J. Sound Vib.* **2015**, *335*, 367–383. [CrossRef]
121. Jiang, L.; Shi, T.; Xuan, J. Fault diagnosis of rolling bearings based on marginal fisher analysis. *J. Vib. Control* **2014**, *20*, 470–480. [CrossRef]
122. Yu, J. Local and global principal component analysis for process monitoring. *J. Process Control* **2012**, *22*, 1358–1373. [CrossRef]
123. Chen, J.; Ma, Z.; Liu, Y. Local coordinates alignment with global preservation for dimensionality reduction. *IEEE Trans. Neural Netw. Learn.* **2013**, *24*, 106–117. [CrossRef]
124. Zhang, M.; Ge, Z.; Song, Z.; Fu, R. Global-local structure analysis model and its application for fault detection and identification. *Ind. Eng. Chem. Res.* **2011**, *50*, 6837–6848. [CrossRef]
125. Gao, Y.; Yu, D.; Wang, H.; Chen, T. Fault feature extraction method of rolling bearing based on spectral graph indices. *J. Aeronaut. Power* **2018**, *33*, 2033–2040.
126. Cheng, J.; Zheng, J.; Yang, Y.; Luo, S. Fault diagnosis model for rolling bearing based on partly ensemble local characteristic-scale decomposition and Laplacian score. *J. Vib. Eng.* **2014**, *27*, 942–950.
127. Vakharia, V.; Gupta, V.; Kankar, P. Efficient fault diagnosis of ball bearing using ReliefF and Random Forest classifier. *J. Braz. Soc. Mech. Sci. Eng.* **2017**, *39*, 2969–2982. [CrossRef]
128. Li, Y.; Yang, Y.; Li, G.; Xu, M.; Huang, W. A fault diagnosis scheme for planetary gearboxes using modified multi-scale symbolic dynamic entropy and mRMR feature selection. *Mech. Syst. Signal Process.* **2017**, *91*, 295–312. [CrossRef]
129. Wang, X.; Qiu, J.; Liu, G. New feature selection method in machine fault diagnosis. *Chin. J. Mech. Eng.* **2005**, *18*, 251–254. [CrossRef]
130. Pan, X.; Huang, J.; Mao, H.; Liu, Z. Fault-characteristic extracting technology based on particle swarm optimization. *J. Vib. Shock* **2008**, *27*, 144–147.
131. Kadri, O.; Mouss, L.; Mouss, M. Fault diagnosis of rotary kiln using SVM and binary ACO. *J. Mech. Sci. Technol.* **2012**, *26*, 601–608. [CrossRef]
132. Zhang, X.; Zhang, Q.; Chen, M.; Sun, Y.; Qin, X.; Li, H. A two-stage feature selection and intelligent fault diagnosis method for rotating machinery using hybrid filter and wrapper method. *Neurocomputing* **2017**, *275*, 2426–2439. [CrossRef]
133. Xue, R.; Zhao, R. The fault feature selection algorithm of combination of ReliefF and QPSO. *J. Vib. Shock* **2020**, *39*, 176–181+213.
134. Zhu, X.; Zhang, Y.; Zhu, Y. Intelligent fault diagnosis of rolling bearing based on kernel neighborhood rough sets and statistical features. *J. Mech Sci. Technol.* **2012**, *26*, 2649–2657. [CrossRef]
135. Zhao, X.; Tang, X.; Zhao, J.; Zhang, Y. Fault diagnosis of asynchronous induction motor based on BP neural network. In Proceedings of the International Conference on Measuring Technology and Mechatronics Automation, Changsha, China, 13–14 March 2010.
136. Gunerkar, R.S.; Jalan, A.K.; Belgamwar, S.U. Fault diagnosis of rolling element bearing based on artificial neural network. *J. Mech. Sci. Technol.* **2019**, *33*, 505–511. [CrossRef]
137. Zhu, K.; Chen, L.; HU, X. Rolling element bearing fault diagnosis based on multi-scale global fuzzy entropy, multiple class feature selection and support vector machine. *Trans. Inst. Meas. Control* **2019**, *41*, 4013–4022. [CrossRef]
138. Chen, X.; He, W.; Ma, D.; Zhao, D. Symbol entropy and svm based rolling bearing fault diagnosis. *China Mech. Eng.* **2010**, *21*, 67–70.
139. Tian, Y.; Ma, J.; Lu, C.; Wang, Z. Rolling bearing fault diagnosis under variable conditions using LMD-SVD and extreme learning machine. *Mech. Mach. Theory* **2015**, *90*, 175–186. [CrossRef]
140. Han, T.; Jiang, D. Rolling bearing fault diagnostic method based on VMD-AR model and random forest classifier. *Shock Vib.* **2016**, *6*, 1–11. [CrossRef]
141. Zhang, X.; Zhang, Q.; Qin, X.; Sun, Y. Rolling bearing fault diagnosis based on ITD Lempel-Ziv complexity and PSO-SVM. *J. Vib. Shock* **2016**, *35*, 102–107+138.
142. Tang, G.; Pang, B.; Tian, T.; Zhou, C. Fault diagnosis of rolling bearings based on improved fast spectral correlation and optimized random forest. *Appl. Sci.* **2018**, *8*, 1859. [CrossRef]
143. Wu, J.; Qin, W.; Liang, H.; Jin, S.; Luo, W. Transformer fault identification method based on self-adaptive extreme learning machine. *Elec. Power Autom. Equip.* **2019**, *39*, 181–186.
144. Zhao, R.; Yan, R.; Chen, Z.; Mao, K.; Wang, P.; Gao, R. Deep learning and its applications to machine health monitoring. *Mech. Syst. Signal Process.* **2019**, *115*, 213–237. [CrossRef]
145. Singh, S.; Howard, C.; Hansen, C. Convolutional neural network based fault detection for rotating machinery. *J. Sound Vib.* **2016**, *377*, 331–345.
146. Wen, L.; Li, X.; Gao, L.; Zhang, Y. A new convolutional neural network-based data-driven fault diagnosis method. *IEEE Trans. Ind. Electron.* **2017**, *65*, 5990–5998. [CrossRef]
147. Wang, J.; Zhuang, J.; Duan, L.; Cheng, W. A multi-scale convolution neural network for featureless fault diagnosis. In Proceedings of the International Symposium on Flexible Automation, Cleveland, OH, USA, 1–3 August 2016.
148. Zhang, W.; Peng, G.; Li, C.; Chen, Y.; Zhang, Z. A new deep learning model for fault diagnosis with good anti-noise and domain adaptation ability on raw vibration signals. *Sensors* **2017**, *17*, 425. [CrossRef]

149. Sun, W.; Deng, A.; Deng, M.; Zhu, J.; Zhai, Y. Multi-view feature fusion for rolling bearing fault diagnosis using random forest and autoencoder. *J. Southeast Univ.* **2019**, *35*, 33–40.
150. Shi, P.; Guo, X.; Han, D.; Fu, R. A sparse auto-encoder method based on compressed sensing and wavelet packet energy entropy for rolling bearing intelligent fault diagnosis. *J. Mech. Sci. Technol.* **2020**, *34*, 1445–1458. [CrossRef]
151. Zhou, X.; Zhang, X.; Zhang, W.; Xia, X. Fault diagnosis of rolling bearing under fluctuating speed and variable load based on TCO spectrum and stacking auto-encoder. *Measurement* **2019**, *138*, 162–174.
152. Gu, Y.; Cao, J.; Song, X.; Yao, J. A Denoising autoencoder-based bearing fault diagnosis system for time-domain vibration signal. *Wirel. Commun. Mob. Com.* **2021**, *2021*, 9790053. [CrossRef]
153. Oh, H.; Jung, J.H.; Jeon, B.C.; Youn, B.D. Scalable and unsupervised feature engineering using vibration-imaging and deep learning for rotor system diagnosis. *IEEE Trans. Ind. Electron.* **2018**, *65*, 3539–3549. [CrossRef]
154. Chen, Z.; Li, W. Multisensor feature fusion for bearing fault diagnosis using sparse autoencoder and deep belief network. *IEEE Trans. Instrum. Meas.* **2017**, *66*, 1693–1702. [CrossRef]
155. Shao, H.; Jiang, H.; Li, X.; Liang, T. Rolling bearing fault detection using continuous deep belief network with locally linear embedding. *Comput. Ind.* **2018**, *96*, 27–39. [CrossRef]
156. Yuan, M.; Wu, Y.; Lin, L. Fault diagnosis and remaining useful life estimation of aero engine using LSTM neural network. In Proceedings of the IEEE International Conference on Aircraft Utility Systems, Austin, TX, USA, 30 October 2016.
157. Zhao, R.; Wang, J.; Yan, R.; Mao, K. Machine health monitoring with LSTM networks. In Proceedings of the International Conference on Sensing Technology, Nanjing, China, 11–13 November 2016.
158. Zhao, K.; Shao, H. Intelligent fault diagnosis of rolling bearing using adaptive deep gated recurrent unit. *Neural Process. Lett.* **2020**, *51*, 1165–1184. [CrossRef]
159. Case Western Reserve University Bearing Data Center. Available online: http://csegroups.case.edu/bearingdatacenter/home/ (accessed on 1 April 2018).
160. Nectoux, P.; Gouriveau, R.; Medjaher, K.; Ramasso, E.; Chebel-Morello, B.; Zerhouni, N.; Varnier, C. PRONOSTIA: An experimental platform for bearings accelerated degradation tests. In Proceedings of the IEEE International Conference on Prognostics and Health Management, PHM'12, Mineapollis, MN, USA, 23–27 September 2012.
161. Gousseau, W.; Antoni, J.; Girardin, F.; Griffaton, J. Analysis of the Rolling Element Bearing data set of the Center for Intelligent Maintenance Systems of the University of Cincinnati. In Proceedings of the CM2016, Paris, France, 10–12 October 2016.
162. Huang, H.; Baddour, N. Bearing vibration data collected under time-varying rotational speed conditions. *Data Brief* **2018**, *21*, 1745–1749. [CrossRef]
163. Wang, B.; Lei, Y.; Li, N.; Li, N. A hybrid prognostics approach for estimating remaining useful life of rolling element bearings. *IEEE Trans. Reliab.* **2020**, *69*, 401–412. [CrossRef]
164. Deng, J.; Dong, W.; Socher, R.; Li, L.; Li, K.; Li, F. ImageNet: A large-scale hierarchical image databas. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20–25 June 2009.
165. Weiss, K.; Khoshgoftaar, T.M.; Wang, D. A survey of transfer learning. *J. Big Data* **2016**, *3*, 1–40. [CrossRef]
166. Guo, L.; Lei, Y.; Xing, S.; Yan, T.; Li, N. Deep convolutional transfer learning network: A new method for intelligent fault diagnosis of machines with unlabeled data. *IEEE Trans. Ind. Electron.* **2018**, *66*, 7316–7325. [CrossRef]
167. Zhang, M.; Wang, D.; Lu, W.; Yang, J.; Li, Z.; Liang, B. A deep transfer model with wasserstein distance guided multi-adversarial networks for bearing fault diagnosis under different working conditions. *IEEE Access* **2019**, *7*, 65303–65318. [CrossRef]
168. Li, X.; Zhang, W.; Ma, H.; Luo, Z.; Li, X. Deep learning-based adversarial multi-classifier optimization for cross-domain machinery fault diagnostics. *J. Manuf. Syst.* **2020**, *55*, 334–347. [CrossRef]
169. Wang, D.; Zhang, M.; Xu, Y.; Lu, W.; Yang, J.; Zhang, T. Metric-based meta-learning model for few-shot fault diagnosis under multiple limited data conditions. *Mech. Syst. Signal Process.* **2021**, *155*, 107510. [CrossRef]
170. Wu, J.; Zhao, Z.; Sun, C.; Yan, R.; Chen, X. Few-shot transfer learning for intelligent fault diagnosis of machine. *Measurement* **2020**, *166*, 108202. [CrossRef]
171. Fuan, W.; Hongkai, J.; Haidong, S.; Wenjing, D.; Shuaipeng, W. An adaptive deep convolutional neural network for rolling bearing fault diagnosis. *Meas. Sci. Technol.* **2017**, *28*, 095005. [CrossRef]
172. Gao, S.; Xu, L.; Zhang, Y.; Pei, Z. Rolling bearing fault diagnosis based on intelligent optimized self-adaptive deep belief network. *Meas. Sci. Technol.* **2020**, *31*, 055009. [CrossRef]
173. Tong, J.; Luo, J.; Pan, H.; Zheng, J.; Zhang, Q. A Novel Cuckoo Search Optimized Deep Auto-Encoder Network-Based Fault Diagnosis Method for Rolling Bearing. *Shock Vib.* **2020**, *2020*, 8891905. [CrossRef]
174. Xiao, M.; Zhang, W.; Wen, K.; Zhu, Y.; Yiliyasi, Y. Fault Diagnosis Based on BP Neural Network Optimized by Beetle Algorithm. *Chin. J. Mech. Eng.* **2021**, *34*, 119. [CrossRef]
175. Chen, J.; Jiang, J.; Guo, X.; Tan, L. A self-Adaptive CNN with PSO for bearing fault diagnosis. *Syst. Sci. Control Eng.* **2021**, *9*, 11–22. [CrossRef]
176. Koza, J. *Genetic Programming: On The Programming of Computers by Means of Natural Selection*; MIT Press: Cambridge, MA, USA, 1992.
177. Bi, Y.; Xue, B.; Zhang, M. *Genetic Programming for Image Classification: An Automated Approach to Feature Learning*; Springer International Publishing: Berlin/Heidelberg, Germany, 2021.

178. Shao, L.; Liu, L.; Li, X. Feature learning for image classification via multiobjective genetic programming. *IEEE Trans. Neural Netw. Learn. Syst.* **2014**, *25*, 1359–1371. [CrossRef]
179. Fu, W.; Johnston, M.; Zhang, M. Genetic programming for edge detection: A Gaussian-based approach. *Soft. Comput.* **2016**, *20*, 1231–1248. [CrossRef]
180. Bi, Y.; Xue, B.; Zhang, M. Genetic programming with image-related operators and a flexible program structure for feature learning in image classification. *IEEE Trans. Evolut. Comput.* **2020**, *25*, 87–101. [CrossRef]
181. Bi, Y.; Xue, B.; Zhang, M. An effective feature learning approach using genetic programming with image descriptors for image classification. *IEEE Comput. Intell. Mag.* **2020**, *15*, 65–77. [CrossRef]
182. Guo, H.; Jack, L.; Nandi, A. Feature generation using genetic programming with application to fault classification. *IEEE Trans. Syst. Man Cyber. Part B* **2005**, *35*, 89–99. [CrossRef]
183. Peng, B.; Wan, S.; Bi, Y.; Xue, B.; Zhang, M. Automatic feature extraction and construction using genetic programming for rotating machinery fault diagnosis. *IEEE Trans. Cyber.* **2020**, *51*, 4909–4923. [CrossRef]
184. Peng, B.; Bi, Y.; Xue, B.; Zhang, M.; Wan, S. Multi-view feature construction using genetic programming for rolling bearing fault diagnosis. *IEEE Comput. Intell. Mag.* **2021**, *16*, 79–94. [CrossRef]