

Current Status of Research on Fault Diagnosis Using Machine Learning for Gear Transmission Systems

Xuezhong Fu ^{1,2,3,*} , Yuanxin Fang ¹, Yingqiang Xu ², Haijun Xu ⁴, Guo Ma ³ and Nanjiang Peng ³

¹ School of Mechanical and Automotive Engineering, Guangxi University of Science and Technology, Liuzhou 545006, China; 20220102008@stdmail.gxust.edu.cn

² School of Mechanical Engineering, Northwestern Polytechnical University, Xi'an 710072, China; xuyngqng@nwpu.edu.cn

³ Liuzhou Wuling Automobile Industry Co., Ltd., Liuzhou 545007, China; maguo@wuling.com.cn (G.M.); pengnanjiang@wuling.com.cn (N.P.)

⁴ Liuzhou Wuling New Energy Automobile Co., Ltd., Liuzhou 545005, China; xuhj@wuling.com.cn

* Correspondence: 100002350@gxust.edu.cn

Abstract: Gear transmission system fault diagnosis is crucial for the reliability and safety of industrial machinery. The combination of mathematical signal processing methods with deep learning technology has become a research hotspot in fault diagnosis. Firstly, the development and status of gear transmission system fault diagnosis are outlined in detail. Secondly, the relevant research results on gear transmission system fault diagnosis are summarized from the perspectives of time-domain, frequency domain, and time-frequency-domain analysis. Thirdly, the relevant research progress in shallow learning and deep learning in the field of fault diagnosis is explained. Finally, future research directions for gear transmission system fault diagnosis are summarized and anticipated in terms of the sparsity of signal analysis results, separation of adjacent feature components, extraction of weak signals, identification of composite faults, multi-factor combinations in fault diagnosis, and multi-source data fusion technology.

Keywords: gear transmission system; fault diagnosis; time-frequency-domain analysis; shallow learning; deep learning



Citation: Fu, X.; Fang, Y.; Xu, Y.; Xu, H.; Ma, G.; Peng, N. Current Status of Research on Fault Diagnosis Using Machine Learning for Gear Transmission Systems. *Machines* **2024**, *12*, 679. <https://doi.org/10.3390/machines12100679>

Academic Editor: Ciro Santus

Received: 3 September 2024

Revised: 24 September 2024

Accepted: 26 September 2024

Published: 27 September 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Gear transmission systems have the advantages of high transmission efficiency, high transmission accuracy, reliable operation, and long service life. They are widely used in automobile manufacturing, aerospace, and other fields. Gears, as the core of gear transmission systems, are often subjected to complex working conditions with alternating loads and are prone to failure. The main failures of gears include wear, bonding, and fracture. According to statistics, gearbox failures account for about 22% of all mechanical component failures [1]. Among all types of gearbox damage, gear failure accounts for 60%. Thus, gear failure is the main cause of mechanical system failure in gear transmission systems. Therefore, studying effective gear transmission system fault diagnosis methods to detect and repair faults in a timely manner has significant practical value for improving the safety and reliability of mechanical operations.

Fault detection and diagnosis methods, such as vibration analysis, oil analysis, temperature analysis, energy analysis, and acoustic emission detection, are the primary means of extracting fault features and diagnosing faults in gear transmission systems. Compared with several other methods, vibration signal analysis offers the advantages of simple measurement and easy analysis. Relevant research is needed. The diagnosis process is shown in Figure 1. In the early stages of gear transmission system failure, its prediction and maintenance are extremely important. By analyzing the tooth surface contact point and transmission error, gear failure can be effectively predicted. Gear transmission errors,

misalignment, and other factors can cause localized contact stress concentrations during gear meshing, leading to cracks or pitting, often accompanied by vibration and noise. [2,3]; gear tooth profile geometric deviation and shaft angle changes will lead to edge contact, increasing the risk of gear fracture and fatigue failure [4]; severe misalignment and stress concentration often accelerate the crack propagation of gears, which in turn leads to system performance degradation and even causes sudden gear fracture [5]. TCA technology plays an important role in this process [6]. It can accurately simulate the theoretical transmission error and contact characteristics of the gear system, analyze the load distribution during gear meshing, and achieve more accurate fault identification and prediction by combining vibration analysis, stress detection, and other technologies. When a gear transmission system fails, the energy distribution and frequency components of its vibration signal change. The fault characteristic signals in the vibration signal can be extracted through signal processing methods. At present, the methods for vibration signal processing are mainly divided into two categories: one is the traditional mathematical processing method represented by the fast Fourier transform [7], which mainly includes time-domain analysis methods, frequency domain statistical analysis methods, Fourier transform analysis methods, time series model analysis methods, wavelet transform methods, empirical mode decomposition methods, local mean decomposition methods, Hilbert-Huang transform methods, second-generation wavelet transform methods, morphological component methods, maximum correlation kurtosis deconvolution methods, etc.; the other is intelligent diagnostic technologies represented by neural networks [8], which mainly include intelligent diagnosis technologies based on expert systems, intelligent diagnosis technologies based on artificial neural networks, intelligent diagnosis technologies based on fuzzy logic, intelligent diagnosis technologies based on genetic algorithms, intelligent diagnosis technologies based on fuzzy neural networks, intelligent diagnosis technologies based on support vector machines, intelligent diagnosis technologies based on convolutional neural networks, and intelligent diagnosis technologies based on deep belief networks. Traditional signal processing methods generally require extensive professional knowledge, a deep understanding of signal characteristics, and substantial practical experience; intelligent diagnosis technologies require a lot of iterative calculations on vibration signals, which have problems of low efficiency and low accuracy.

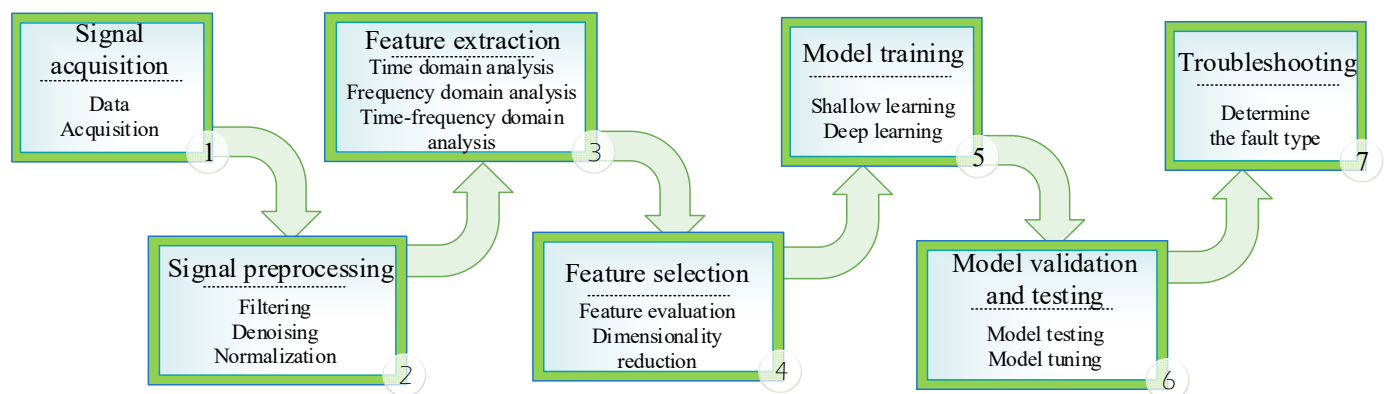


Figure 1. Gear transmission system fault diagnosis flow chart.

This paper reviews some diagnostic methods that are currently widely used in gear transmission system fault diagnosis from the perspective of vibration signal analysis, further organizes the latest research results in gear transmission system fault diagnosis, and finally, looks forward to future research directions in gear transmission system fault diagnosis, providing a reference for the next stage of research.

2. Development of Fault Diagnosis of Gear Transmission System

As early as the beginning of the 20th century, researchers began to study the vibration and noise generated by gears. By the 1960s, vibration and noise were identified as impor-

tant indicators for evaluating the performance of gear transmission systems [9]. In 1968, Optiz [10] proposed that the vibration and noise of gears are functions of gear transmission power, transmission error, and accuracy. In the 1970s, the Fourier transform became an important tool. Randall et al. [11] used the dynamic characteristics of gearbox components to detect the vibration signals of gears and analyze gearbox faults, achieving remarkable results in analyzing gear wear and cracks. In the 1980s, the advancement of mathematical signal processing technology made vibration signal analysis more accurate and diversified. Time-frequency analysis methods (such as short-time Fourier transform wavelet transform) and envelope analysis technology have been widely used in gear fault diagnosis and have proven effective in processing non-stationary signals. McFadden et al. [12] analyzed the time-frequency distribution of gear vibration signals through spectrum diagrams and applied image processing technology for fault detection. Forrester [13] proposed using the Wigner-Ville distribution to analyze vibration signals for early gear fault detection, effectively identifying the type and degree of gear faults. In the 1990s, CAT technology was gradually applied to the field of gear fault diagnosis. Yu et al. [14] developed a gear noise CAT test system based on the Fourier transform, which met the requirements of real-time testing in industrial production. With the rapid development of computer technology and artificial intelligence, intelligent fault diagnosis technologies, such as neural networks, support vector machines, and deep learning, have become widely used in gear fault diagnosis. Samanta [15] used genetic algorithms to optimize the characteristics of gear vibration signals and input them into artificial neural networks and support vector machines for classification. In recent years, scholars have focused on combining mathematical signal analysis technology with intelligent diagnosis to diagnose gear transmission system faults, thereby significantly improving the efficiency and accuracy of fault diagnosis. Cheng et al. [16] proposed a method of combining Hilbert-Huang transform with self-organizing feature mapping neural networks to identify gear faults. Hilbert-Huang transform effectively processes nonlinear and non-stationary signals. Combined with the adaptive and self-learning characteristics of the self-organizing map neural network, this method can accurately extract and distinguish gear fault characteristics. Sanz et al. [17] combined discrete wavelet transforms with multi-layer perceptron neural networks to achieve gear fault detection and location. Hu et al. [18] used discrete wavelet decomposition and an improved convolutional neural network to achieve efficient fault identification under limited sample data. In summary, fault diagnosis technology for gear transmission systems has evolved from traditional mathematical signal processing to intelligent diagnosis technology. In recent years, the development trend in fault diagnosis technology has been to combine intelligent algorithms with advanced signal processing technology to achieve more efficient and accurate fault pattern recognition and classification.

3. Fault Diagnosis Method Based on Mathematical Signal Processing

3.1. Time-Domain Analysis Method

Time-domain analysis [19] processes signals in the time-domain according to a certain time sequence. It is characterized by strong sensitivity, a high signal-to-noise ratio, and the ability to effectively highlight the characteristic information of gear faults. The diagnostic process is shown in Figure 2. However, it can only qualitatively analyze possible faults in the gear. For confirmed faults, it needs to be combined with other methods. Therefore, this method is suitable for early gear fault detection.

Time-domain synchronous averaging is a typical time-domain analysis method that can effectively extract periodic components from complex signals, remove irrelevant components, and improve signal-to-noise ratio. It is especially suitable for gear fault diagnosis with obvious periodicity. In 1987, Bond [20] first introduced the synchronous averaging method to reduce periodic noise in vibration signals. To address the limitations of the synchronous averaging method in processing non-periodic and variable-speed signals, Isermann [21] proposed the weighted synchronous averaging method, which enhances the detection capability of specific fault features by adjusting the weights. Fu et al. [22]

proposed a multi-channel synchronous averaging technology to improve the accuracy of gearbox fault diagnosis. The above research methods are all based on time-domain signal analysis and processing. However, the time-domain synchronous averaging method may have difficulty capturing the frequency-domain changes of the signal when processing non-stationary signals. The instantaneous frequency change is crucial for detecting fault characteristics. The time-frequency-domain synchronous averaging method combines time-domain and frequency-domain information, allowing for more effective processing of non-stationary signals. Zhang et al. [23] proposed a time-domain synchronous averaging method that does not require a key phase signal. It estimates the instantaneous speed through time-frequency-domain filtering and Hilbert transform, effectively eliminating the influence of speed fluctuations and monitoring gear faults. Ha et al. [24] proposed a fault diagnosis method based on the combination of an autocorrelation function and windowed time-domain synchronous averaging. The shape and range of the window function are defined by an autocorrelation function. It has efficient signal processing capabilities and prevents signal distortion during the processing process. It can effectively cope with the status monitoring of planetary gearboxes. With the rapid development of artificial intelligence technology, the shortcomings of traditional methods for processing complex data have become increasingly obvious. Artificial intelligence technology can quickly process large-scale data and automatically extract and analyze complex features. Yao et al. [25] proposed combining synchronous averaging with machine learning technology to improve the intelligence level of gear fault diagnosis. The synchronous averaging method is mainly used for processing periodic signals. By averaging multiple periodic signals, it can effectively reduce the impact of random noise. However, its processing capabilities are weak when facing nonlinear and non-stationary signals. Maximum correlation kurtosis deconvolution (MCKD) is similar to the synchronous averaging method. It maximizes the correlation kurtosis of the input signal under different given periods, enabling the extraction of weak fault signal features from strong background noise. In 2007, Zhang et al. [26] introduced the maximum correlation analysis method based on kurtosis and discussed its application in gear fault diagnosis. However, when dealing with non-stationary signals or complex background noise, the aliasing effect of the maximum correlation kurtosis analysis makes it difficult to extract gear fault features. Liu et al. [27] proposed a method combining deconvolution technology to improve the effect of signal separation and feature extraction. Given the shortcomings of the initial deconvolution technology in dealing with complex noise and nonlinear systems, Wu et al. [28] proposed an improved maximum correlation kurtosis deconvolution method to address the challenges of complex noise and signal nonlinearity. To address the problem of weak signals and strong background noise, Hua et al. [29] proposed a method based on MCKD and multiwavelet decomposition. The MCKD processing of the components of multiwavelet decomposition can effectively suppress background noise and enhance the weak impact characteristics of the signal. Its effectiveness was verified through experimental analysis and engineering applications. For complex vibration signals, the time-domain analysis method has limitations in feature extraction and lacks strong anti-interference ability. Chen et al. [30] combined multi-scale and time-frequency-domain analysis techniques to improve the signal feature extraction capability of MCKD. He et al. [31] proposed a method that combines resonant sparse decomposition with MCKD to diagnose gearbox faults through envelope spectrum analysis. Hong [32] successfully combined an MCKD with multi-scale wavelets to enable the diagnosis of composite faults in aircraft engine rotors. He et al. [33] explored a combination of machine learning and MCKD to improve the complex data processing and fault diagnosis capabilities.

To sum up, the time-domain analysis method can directly reflect changes in vibration signals through time series data and is effective for processing non-stationary signals. However, for complex vibration signals, it is difficult to extract useful fault features directly from time-domain data, and its anti-interference ability is poor. Table 1 compares the advantages, disadvantages, and usage scenarios of various signal analysis methods based

on time-domain analysis. In this context, many scholars have conducted extensive research and proposed methods that combine multi-scale analysis to improve the accuracy and robustness of fault feature extraction, combined with other signal processing methods, such as frequency-domain analysis and time-frequency analysis, to enhance the accuracy of fault diagnosis. In addition, deep learning algorithms are increasingly being used to perform feature extraction and pattern recognition on time-domain data, significantly enhancing their ability to perform automated fault diagnosis.

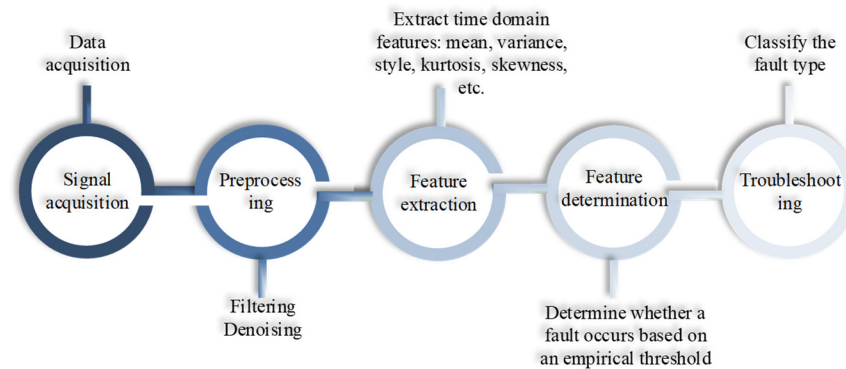


Figure 2. Diagnosis flow chart of the time-domain analysis method.

Table 1. Comparison of the advantages and disadvantages of various analysis methods based on time-domain analysis.

Method	Advantage	Disadvantages	Applicable Scenarios
Time-domain statistics analysis	The calculation is simple, easy to understand and implement	Insensitive to non-Gaussian signals and susceptible to noise interference	Early fault detection, low noise environment
Time-domain waveform analysis	Intuitive, find obvious fault characteristics	Relying on experience and observation, it is difficult to analyze complex signals	Obvious fault feature detection, simple gear transmission system
Envelope analysis	Effectively extract impact signals suitable for fault feature extraction	Requires appropriate filter design, high computational complexity	Early gear fault detection
Shock pulse method	Sensitive to early faults and capture short-term impact signals	Requires high sampling rate equipment, noise sensitive	Early fault diagnosis of gears, local defects on tooth surface
Time-domain synchronous averaging	Suppress random noise and improve signal-to-noise ratio, especially suitable for extracting periodic impact signals	Requires precise synchronization signal, sensitive to speed fluctuations, and has complex processing	Gearbox periodic fault diagnosis, noisy environment
Maximum correlation kurtosis deconvolution	Significantly enhances impact signals, is sensitive to early faults, and extracts weak fault signals	The calculation complexity is high, parameter selection requires experience, and adaptability to strong noise environments needs to be studied.	Early fault diagnosis, weak impact signal extraction

3.2. Frequency-Domain Analysis Method

The frequency-domain analysis method reveals the frequency components of the signal by converting the time-domain signal into the frequency-domain. The Fourier transform is usually used to extract the characteristics of the time-domain waveform and phase waveform. The diagnosis process is shown in Figure 3. This method can clearly identify periodic and steady-state fault characteristics, which is better than time-domain analysis. Therefore, it is widely used in fault diagnosis.

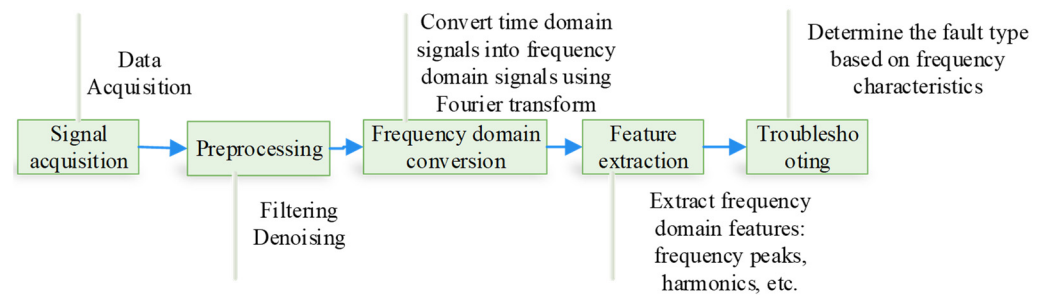


Figure 3. Diagnosis flow chart of the frequency-domain analysis method.

In 1822, Fourier laid the foundation of the Fourier transform theory in his classic work. However, the classical Fourier transform (CTFT) can only provide frequency-domain information and cannot describe the characteristics of the signal changing over time. It has limitations in analyzing transient or non-stationary signals. For this reason, Gabor proposed the short-time Fourier transform (STFT), which introduces a sliding window and divides the signal into segments for the Fourier transform, enabling it to process non-stationary signals. STFT has been widely used in the time-frequency analysis of non-stationary vibration signals and combined with other algorithms for fault diagnosis. Wang et al. [34] took advantage of the STFT to process non-stationary signals and proposed a gear fault monitoring method based on electrostatic signals and STFT, extracting fault features through time-frequency analysis. Wu et al. [35] proposed an optimization algorithm based on short-time Fourier transform and synchronous compression theory, which improved the time-frequency resolution and enhanced the energy concentration level of the signal. Bao et al. [36] proposed an adaptive short-time Fourier transform method, which adaptively adjusted the time window length through fast path optimization to obtain better time-frequency resolution. Based on the traditional Fourier transform, to make it have a wider signal representation capability, the fractional Fourier transform (FRFT) was proposed. By introducing a fractional order parameter, the signal can be rotated at any angle between the time-domain and the frequency-domain, which can more effectively process certain specific types of non-stationary signals, such as linear frequency modulation signals and modulated signals. Zhang and Jiao [37,38] conducted a detailed analysis of the principle and application of the fractional Fourier transform, pointing out that it is extremely suitable for processing time-varying non-stationary signals. Hu et al. [39] proposed an optimal window function Gabor transform method based on the fractional Fourier transform, which improved the fault diagnosis capability of planetary gearboxes under strong noise background. In addition, many scholars have conducted research on the superiority of fractional Fourier transform in processing linearly correlated signals and modulated signals [40–42]. In recent years, many scholars have conducted research on combining the Fourier transform with deep learning technology. Guo et al. [43] proposed a gear fault diagnosis method based on meta-learning technology to solve the problems of a small number of fault samples and large differences in data distribution in the process of gear fault diagnosis under variable working conditions. The original vibration signal is overlapped, and a short-time Fourier transform is performed to make its data form more consistent with the input of the model. Finally, the gear fault is diagnosed using meta-learning technology. Yu et al. [44] proposed a gearbox intelligent fault diagnosis method based on a short-time Fourier transform and convolutional neural network to solve the problem of low fault recognition rate of shallow machine learning methods. The gear vibration signal is transformed by STFT to obtain a time-frequency diagram and input it into the CNN model for fault diagnosis. In addition, many scholars [45–48] have conducted research on using an improved Fourier transform algorithm to process fault signals and automatically analyze and classify time-frequency diagrams through deep learning models.

To sum up, the frequency-domain analysis method is widely used in gear fault diagnosis. The advantages, disadvantages, and usage scenarios of each analysis method based on frequency-domain analysis are compared and analyzed, as shown in Table 2. The research

mainly focuses on improving the optimization algorithm, combined with time-frequency analysis, fusion of multi-source information, and deep learning technology. Although frequency domain analysis has problems such as noise sensitivity and difficulty in processing nonlinear and non-stationary signals, research shows that optimization algorithms and combining time-domain and frequency domain analyses can significantly improve the accuracy of fault diagnosis.

Table 2. Comparison of the advantages and disadvantages of various analysis methods based on frequency-domain analysis.

Method	Advantage	Disadvantages	Applicable Scenarios
Envelope demodulation analysis	Effectively extract and amplify impact signals, sensitive to early faults	The filter design is complex, parameter selection requires experience, and it is sensitive to strong noise environments.	Early fault detection, gear fault with obvious impact signal
Resonance demodulation analysis	Enhance weak fault signals, improve signal-to-noise ratio, and be sensitive to early faults	Requires accurate resonant frequency, complex filter design	Weak fault signal detection, tooth surface cracks, and wear
STFT	Provides time-frequency information, suitable for non-stationary signals	The time-frequency resolution is limited, and the window length selection needs to be balanced	Detection of non-stationary signals
Fast Fourier Transform	Fast calculation speed, suitable for real-time processing	Insensitive to non-stationary signals and difficult to process transient signals	Suitable for stable working state, real-time monitoring, and fault diagnosis
FRFT	Sensitive to linear frequency modulation signals, provides time-domain and frequency-domain information, suitable for complex signals	The calculation complexity is high, the selection of fractional order parameters requires experience, and the interpretation of results is complicated	Complex and non-stationary signal detection

3.3. Time-Frequency-Domain Analysis Method

The time-frequency analysis method can map the one-dimensional time-domain signal to the two-dimensional time-frequency-domain, thereby obtaining the time-frequency joint distribution function of the signal. The time-frequency distribution can effectively characterize the time-varying frequency modulation characteristics of the non-stationary signal. The diagnostic process is shown in Figure 4. The time-frequency analysis method effectively processes non-stationary signals and identifies instantaneous fault characteristics by providing time and frequency information at the same time, but the calculation is complex and sensitive to noise. It is suitable for a comprehensive analysis of dynamically changing vibration signals and complex faults.

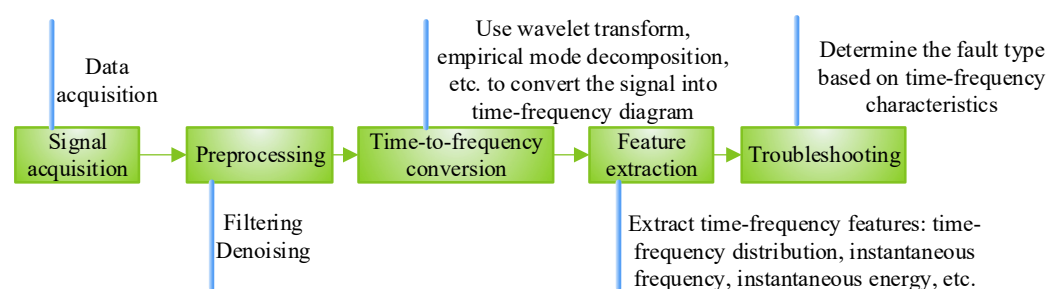


Figure 4. Diagnostic flow chart of the time-frequency analysis method.

Wavelet transform is a typical time-frequency-domain analysis method. It was first proposed in 1974 [49] and has been widely used in many fields, such as signal processing, image analysis, and speech processing. Wavelet transform is based on the Fourier transform

and adopts a localization concept like short-time Fourier transform, which solves the problem of short-time Fourier transform not being able to change the window size as needed. In order to solve the wavelet basis selection, resolution, and calculation problems of the wavelet transform, Mallat et al. [50,51] and many other scholars have proposed optimization methods such as continuous wavelet transform, discrete wavelet transform, wavelet packet transform, and adaptive wavelet transform. In recent years, wavelet transform has achieved significant research results in optimization algorithms and multi-resolution analysis technology combined with other signal processing technologies. Parey et al. [52] proposed a method for analyzing acoustic signals collected under various fault conditions of gearboxes. The acoustic signals obtained under different faults are converted into angular domain signals and then continuously transformed by wavelet transform to obtain the wavelet coefficients of the optimal scale and input them into ANFIS for fault diagnosis. The results show that acoustic signals and adaptive neural fuzzy inference systems can be effectively used for gearbox condition diagnosis. With the development of artificial intelligence, the method of integrating wavelet transform with intelligent diagnosis technology is currently a research hotspot. Liang et al. [53] used wavelet transform as a preprocessing and feature extraction tool combined with a multi-label convolutional neural network to realize intelligent diagnosis of gearbox compound faults. Saravanan et al. [54] used discrete wavelet transform to extract the features of different types of gear fault signals and used decision trees for feature selection and classification, successfully realizing the classification and identification of various faults of spur bevel gears. Wavelet transform has the advantages of processing non-stationary signals, providing good time-frequency localization and multi-resolution analysis, and can effectively identify and locate fault features. However, its computational complexity is high and its sensitivity to the selection of the wavelet basis may affect its application. Wavelet transform is suitable for early fault detection and signal processing in noisy environments. In recent years, the integration of wavelet transform with deep learning and intelligent diagnosis technology has become a research hotspot, further expanding its application prospects.

Empirical mode decomposition (EMD) is an adaptive signal time-frequency processing method that can decompose any modulated signal into several intrinsic mode functions and can more naturally process complex nonlinear and non-stationary signals, overcoming the sensitivity of wavelet transform to the selection of wavelet basis and the limitations of predefined basis functions. In 1998, Huang et al. [55] first proposed an empirical mode decomposition method to process nonlinear and non-stationary signals. However, when processing high-noise signals, there are problems with unstable decomposition results and mode aliasing. Wu et al. [56] proposed ensemble empirical mode decomposition (EEMD) to improve the decomposition results of EMD by introducing noise, reducing the influence of mode aliasing and noise. Zhang et al. [57] proposed a method based on frequency modulation empirical mode decomposition (FM-EMD) to address the problem of modal aliasing in dense frequency conditions. This method is used to process weak nonlinear signals in dense frequency conditions and improve the modal aliasing problem. By combining EMD with other algorithms, the influence of random noise and local strong interference can be effectively handled. Park et al. [58] proposed a method to apply EEMD to transmission errors, which can more quickly classify the tooth spalling and crack faults of faulty gears. Tang et al. [59] proposed a method based on morphological singular value decomposition and EMD. The original vibration signal was reconstructed in the phase space and subjected to singular value decomposition for noise reduction and morphological filtering. The filtered signal was then subjected to empirical mode decomposition to effectively filter out random noise signals. Yang et al. [60] proposed a method that combines EMD and kernel-independent component analysis and further improved the performance through a particle swarm optimization algorithm. They successfully solved the problem of underdetermined blind source separation of nonlinear mixed signals in wind turbine gearbox fault detection. In recent years, EMD combined with deep learning and other technologies has shown great advantages in processing large-scale data and complex patterns. Zhang et al. [61]

explored the method of combining deep learning with EMD to improve the diagnostic ability of complex signals and faults. Empirical mode decomposition has the advantages of adaptive signal decomposition and the absence of predefined basis functions. It shows great advantages in processing nonlinear and non-stationary complex signals. However, under interference conditions, such as strong noise, EMD is prone to modal aliasing problems. To solve this problem, scholars have gradually improved the EMD method and combined it with other algorithms for optimization, thereby improving its robustness to noise. In recent years, with the development of artificial intelligence, combining EMD with machine learning technology has become a research hotspot, which can significantly improve the accuracy and efficiency of fault diagnosis and show broad application prospects.

Like EMD, local mean decomposition (LMD) is a new adaptive processing method for nonlinear and non-stationary signals, as proposed by Jonathan. It decomposes nonlinear and non-stationary signals into multiple product function (PF) components; each PF component is a single-component amplitude-frequency modulated signal, making it suitable for processing non-stationary and nonlinear signals. In 2009, Huang et al. [62] introduced the theoretical basis of basic LMD, emphasizing its application in processing nonlinear and non-stationary signals, and pointed out that there are limitations in processing complex and non-stationary signals, especially the problems of modal aliasing and signal component overlap that may occur during the decomposition process. To address this problem, Niu et al. [63] introduced noise suppression and signal reconstruction technology into LMD and proposed an enhanced local mean decomposition to improve the processing ability of noise and complex signals. To address the problem that the traditional LMD method may not be flexible enough when processing signals with dynamic characteristics, Cheng et al. [64] creatively proposed a local characteristic scale decomposition method based on intrinsic time scale decomposition. This method has the advantages of a short calculation time and no obvious endpoint effect, while ensuring the clear physical meaning of the decomposed components, and has opened new ideas for adaptive time-frequency analysis methods. Cheng et al. [65] proposed a rotating machinery fault diagnosis method based on LMD based on the modulation characteristics of gear fault vibration signals. First, LMD was compared with EMD to demonstrate its superiority. The effectiveness of the method was verified by analyzing the actual gearbox vibration signal. Although the improved LMD method performs well in many cases, it has shortcomings when dealing with strong noise, pseudo-modes, and multi-scale signals. Pan et al. [66] proposed an improved integrated local mean decomposition method to improve the processing capabilities of noise and pseudo-modes by combining integration technology. Zhou et al. [67] introduced a multi-scale local mean decomposition method to improve the analysis capabilities of complex signals. Cui et al. [68] proposed a local mean demodulation method for the multi-component amplitude modulation-frequency modulation characteristics of gear fault vibration signals and combined it with local feature scale decomposition for gear fault diagnosis. When dealing with large-scale data and complex patterns, the traditional LMD method may not be sufficient. Combining deep learning technology can improve the analysis and feature extraction capabilities. Goyal et al. [69] explored the method of combining deep learning with LMD to improve the diagnostic capabilities of complex signals and faults. Local mean decomposition effectively extracts the characteristics of nonlinear and non-stationary signals by decomposing the signal into local mean and oscillation components. It is particularly suitable for the analysis of complex and weak fault signals. However, the calculation process is relatively complex and sensitive to noise. In dealing with the problem of modal aliasing when processing complex signals, LMD has undergone continuous evolution from basic LMD to various improved methods, including adaptive LMD, integrated LMD, and multi-scale LMD. In recent years, methods combined with deep learning have become a research hotspot. These methods can not only automatically extract complex features and improve the accuracy of fault feature recognition, but also enhance the robustness and intelligence level of the system, thereby achieving more efficient and accurate gear fault diagnosis.

In summary, research on time-frequency analysis methods mainly focuses on improving computing efficiency, enhancing noise suppression capabilities, realizing automated feature extraction, and multi-source data fusion. The advantages, disadvantages, and uses of various analysis methods based on time-frequency-domain analysis are also discussed. Scenarios are compared and analyzed, as shown in Table 3. Although most of the research on combining time-frequency analysis with deep learning technology is still in the theoretical exploration stage, research in this direction has significantly broadened the application scope of time-frequency analysis. Research has proven that, compared with traditional time-frequency analysis methods, combining multi-source data fusion and deep learning technology can provide higher diagnostic accuracy and robustness. Therefore, in-depth research on this is important. However, the effective combination of multi-source data technology still faces many challenges, and future research still needs to continue to explore and develop.

Table 3. Comparison of advantages and disadvantages of various analysis methods based on time-frequency-domain analysis.

Method	Advantage	Disadvantages	Applicable Scenarios
Wavelet transform	Sensitive to non-stationary signals and transient faults, strong in multi-scale analysis capabilities	The computational complexity is high, the selection of wavelet basis and scale requires experience, and the interpretation of results is complex	Non-stationary signals and transient fault detection
Hilbert-Huang Transform	Applicable to nonlinear and non-stationary signals, providing high-precision time-frequency information	High computational complexity, susceptible to noise, and complex interpretation of results	Complex fault signal detection and analysis
EMD	Strong adaptability, suitable for nonlinear and non-stationary signals	Susceptible to noise and boundary effects, the decomposition process may produce mode aliasing	Analysis of complex signals, such as various types of gear faults
LMD	Effectively handle non-stationary signals, sensitive to transient faults, and the decomposition results have good physical meaning	High computational complexity, susceptible to noise, and interpretation of results requires experience	Non-stationary signals and transient fault detection, such as shock and sudden faults in gear systems

4. Fault Diagnosis Method Based on Intelligent Diagnosis

Fault diagnosis methods based on intelligent diagnosis have the advantages of automation and intelligence. They are an effective means to build highly robust models and process extremely complex data, and are an important research direction in fault diagnosis. The following will sort out and summarize the commonly used intelligent diagnosis methods, focusing on four main machine learning models: artificial neural networks (ANN), support vector machines (SVM) based on shallow learning, and convolutional neural networks (CNN) and deep belief networks (DBN) based on deep learning.

4.1. Diagnostic Methods Based on Shallow Learning

Shallow learning is a traditional machine learning method that relies on manually designed features and performs fault classification or regression through algorithms such as artificial neural networks, support vector machines, decision trees, or random forests. The diagnostic process is shown in Figure 5. It performs well when the amount of data is limited, and the features are clear, with fast training speed and strong model interpretability, and is suitable for scenarios with small-scale data sets and low computing resources.

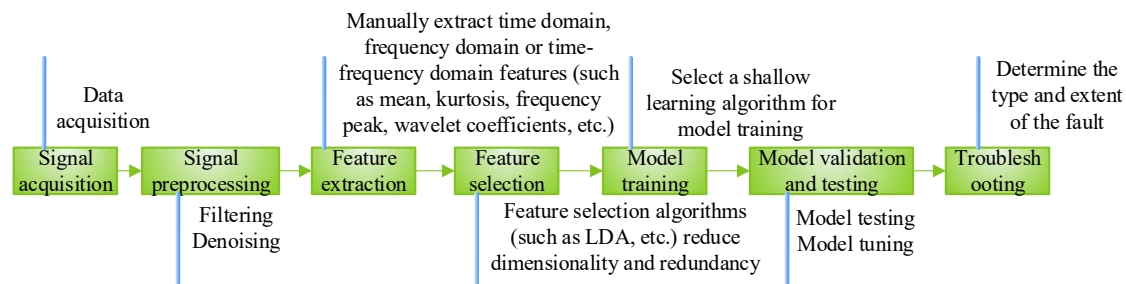


Figure 5. Diagnosis flowchart based on shallow learning.

Artificial neural network models have strong generalization and self-learning capabilities and mainly include back propagation (BP) neural networks, Hopfield neural networks, and Elman neural networks. The BP neural network model has excellent self-learning ability and can approximate any complex nonlinear function. In 1990, Purdue University used BP neural networks to diagnose faults in chemical cracking equipment [70], pointing out the superiority of neural networks in the field of fault diagnosis and promoting the development of neural networks in the field of fault diagnosis. However, the BP network model requires a large amount of sample data, has a slow learning speed, and uses the gradient descent method, which can easily fall into the local optimum. Therefore, Yu [71] introduced genetic algorithms into a BP neural network to optimize the samples and applied the model in practice, showing excellent results. The Hopfield network model has an associative memory function and can introduce an energy function to ensure local minimum convergence and grid stability. In 1982, Hopfield [72] proposed a Hopfield network model that combines neural dynamics with neural networks. In recent years, it has been applied in the field of mechanical fault diagnosis. Jin et al. [73] proposed a wireless sensor fault diagnosis method based on fuzzy theory and discrete Hopfield neural network, which can quickly and effectively diagnose the fault of abnormal sensors and improve the operating reliability of wireless sensors. Skowron et al. [74] used the Hopfield neural network to realize the automatic detection of early faults of stators and rotors in induction motors. The Hopfield network model has the problem of a limited number of memory patterns, and it is difficult to correctly distinguish when the memory patterns are similar. To address this problem, Lang [75] used the status of relays and switches in the ship power grid as fault evaluation factors and performed fault analysis through the Hopfield neural network for the first time. The status of the relays and switches was obtained with the help of fault trees, which were used as training samples of the network to improve the network recognition accuracy. Then, the fault was diagnosed based on the good associative memory ability of the Hopfield neural network, and the practicality of the model was confirmed by simulation experiments. In the 1990s, Elman proposed the Elman neural network model, which has short-term memory, good effect on processing dynamic information, good network stability, and strong computing power, and has excellent application prospects in pattern recognition and other fields. Since the Elman neural network model is prone to fall into local optimality and has weak generalization ability, Lin et al. [76] introduced the artificial bee colony algorithm into the Elman neural network, compared the local optimal solutions obtained by the artificial bee colony, and realized global rapid optimization. In summary, artificial neural networks can handle complex nonlinear relationships and have strong self-learning ability and robustness, but their generalization ability is weak, and they are prone to fall into local optimality. They need to be optimized in combination with other algorithms, and the training process needs to rely on many learning samples.

In the late 1990s, support vector machine network models began to be applied in the field of fault diagnosis [77]. The advantage of the support vector machine model is that it has a strong generalization ability and can obtain global optimality, solving the problem that artificial neural networks are prone to fall into local optimality. In recent years, researchers have combined mature data analysis theories with SVM and further optimized the kernel function of the SVM. Tang et al. [78] used the Shannon wavelet function as the

kernel function of SVM and successfully applied it to the fault diagnosis of wind turbine conversion systems. Yang et al. [79] optimized the kernel parameters of SVM through the artificial bee colony algorithm, obtained the optimal parameter configuration, and successfully applied it to the fault diagnosis of gearboxes. At present, most research is based on a large amount of sample data. In actual application scenarios, the number of fault samples is small, and it takes time to accumulate. Therefore, research on small samples is becoming increasingly important. Martin et al. [80] used resampling technology to synthesize virtual samples to enhance small samples, thereby achieving a balance in sample size and realizing SVM fault diagnosis for induction motors. Ren et al. [81] proposed an improved SVM based on its ability to effectively process small sample data and used principal component analysis and genetic algorithm to optimize parameters. The experimental results verified that fault diagnosis showed better feasibility. Support vector machines have strong classification capabilities for complex problems and good generalization capabilities for small sample data. However, they take a long time to train on large data sets, require large computing resources, and make it difficult to select kernel functions and parameters. They are also not suitable for multi-classification problems. In response to these problems, many scholars have optimized them through optimization algorithms and combined them with other algorithms, but further research is still needed.

In summary, shallow learning methods have lower computational complexity and faster training speed and are easier to explain and implement. However, these methods usually rely on manually designed features and may show limited learning ability when processing complex patterns. Therefore, they are more suitable for fault diagnosis tasks with clear features and small data volumes. The advantages and disadvantages of various fault diagnosis methods based on shallow learning and their usage scenarios are compared and analyzed, as shown in Table 4. Future research directions should include combining deep learning technology with shallow learning methods to achieve automated feature extraction, improve the ability to identify complex and nonlinear fault patterns, and optimize algorithms to cope with larger and more complex data sets.

Table 4. Comparison of the advantages and disadvantages of various fault diagnosis methods based on shallow learning.

Method	Advantage	Disadvantages	Applicable Scenarios
ANN	Strong nonlinear mapping capability, suitable for processing complex high-dimensional data and multi-classification problems	The training is complex, the computation is large, it is easy to fall into the local optimal solution, and a large amount of training data is required	Large-scale data, complex nonlinear problems, such as image classification, etc.
SVM	High-dimensional space performs well, is suitable for small sample data, and has a strong generalization ability	Inefficient for large-scale data and sensitive to parameter selection	Small sample, high-dimensional feature data, binary classification problem
K-nearest neighbor	Simple and intuitive, no training process required, suitable for multi-classification problems	High computational complexity, not applicable to large-scale data, sensitive to noise and unbalanced data	Datasets with smaller sample sizes and fewer features
Decision Tree	Easy to understand and interpret, handles multi-classification problems and nonlinear relationships	Prone to overfitting, sensitive to noise and changes in data distribution	Datasets with clear structures and complex feature relationships
Linear discriminant analysis	Suitable for high-dimensional data dimensionality reduction and classification, high computational efficiency, simple model	Assuming data is normally distributed and the covariance matrix is the same, poor performance for nonlinear data	Linearly separable data, dimensionality reduction, and classification combined application

4.2. Diagnostic Methods Based on Deep Learning

Deep learning is a technology based on neural networks that can automatically extract features from data. It is particularly suitable for dealing with complex nonlinear problems, such as multimodal fault diagnosis of gear transmission systems, and its diagnostic process is shown in Figure 6. It performs well in large-scale data and complex tasks, but requires a lot of computing resources and data, and the model has low interpretability.

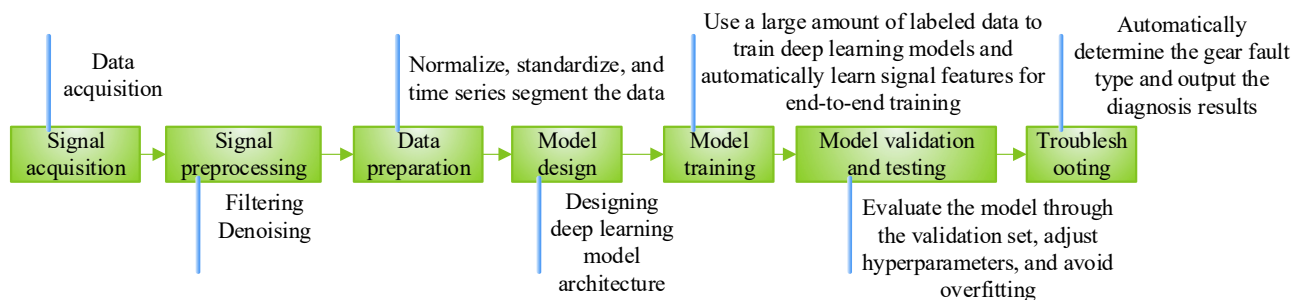


Figure 6. Diagnosis flowchart based on deep learning.

Convolutional neural networks were developed based on inspiration from the cat's visual cortex [82]. Their scaling invariance and local learning characteristics have significant advantages in the field of fault diagnosis. They have high accuracy in diagnosing engineering equipment faults with massive amounts of data, and data features can be automatically extracted and identified. The traditional CNN model is shown in Figure 7. To accurately identify the health status of mechanical equipment, a deep learning model based on a convolutional neural network can be used to process large-capacity, diverse, and high-rate operating data. Dong et al. [83] proposed a method for gearbox health status identification based on a deep neural network, combining variational mode decomposition and wavelet threshold method to reduce the noise of gearbox vibration signal. The experimental results show that the recognition accuracy can exceed 97.5%. Zhang et al. [84] proposed a diagnosis method using a convolutional neural network to process multi-channel fusion data to effectively improve fault identification accuracy, but further research is needed under variable working conditions. Huang et al. [85] established a stable, intelligent fault diagnosis model for planetary gears based on an integrated convolutional neural network to solve the problem of difficulty in identifying planetary gear fault signals and slow convergence speed. The experimental results show that it has a stronger recognition ability and faster convergence speed. CNN has good image data analysis capabilities. When a CNN was introduced into the field of mechanical fault diagnosis, many researchers trained the network by constructing a two-dimensional training set to extract the features of the two-dimensional data and complete the diagnosis. Zhou et al. [86] proposed a new method based on nonlinear autoregressive neural networks and CNN to solve the problem of a small number of fault samples and unbalanced data in practice. The method used a nonlinear autoregressive neural network to expand the fault samples and then used a continuous wavelet transform to convert the one-dimensional vibration signal into a two-dimensional time-frequency image. Finally, a CNN was used for feature learning and fault identification. Liang et al. [87] used wavelet transform to extract two-dimensional time-frequency features from the original one-dimensional vibration signal and input these features into the CNN model to effectively realize the diagnosis of gearbox faults. In summary, convolutional neural networks have shown significant advantages in processing many training samples and image data analysis. It can automatically extract multi-level feature information and has good spatial invariance and position insensitivity in image data processing. The convolution operation effectively reduces the number of parameters and improves the computational efficiency. However, the design and debugging of CNN structures are complex, and the training and inference processes require a lot of computing resources.

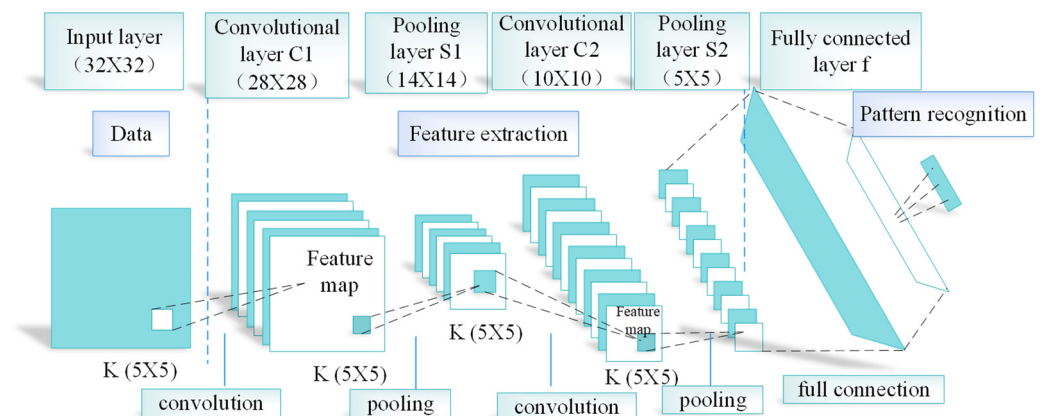


Figure 7. Schematic diagram of the convolutional neural network structure.

Compared with convolutional neural networks, deep belief networks can effectively process high-dimensional and complex data by providing a layer-by-layer pre-training method, solving the gradient vanishing and training difficulties that convolutional neural networks may encounter during training. The structural diagram is shown in Figure 8. In 2013, Tamilselvan et al. [88] first applied the deep belief network model to the field of fault diagnosis and proved its effectiveness by diagnosing the health of aircraft engines and power transformers. On this basis, Yin et al. [89] proposed a DBN-based high-speed train onboard equipment fault diagnosis method, using a DBN to automatically extract features and classify them in the fault frequency-domain spectrum. Compared with traditional neural networks, the classification accuracy is significantly improved. The deep belief network model has good compatibility with other algorithms and can automatically extract features. Zhang et al. [90] further improved the feature extraction performance of the network by optimizing the learning rate and momentum factor of DBN and combined with the multi-core support vector machine optimized by the particle swarm algorithm to realize the fault diagnosis of the transformer. Yan et al. [91] proposed a deep belief neural network based on a genetic algorithm to adaptively adjust parameters and used gas turbines as fault diagnosis objects. The model was compared with the BP, RBF, and SVM models. The results showed that the average diagnostic accuracy of the DBN algorithm was significantly better than that of the other models. In view of the limitation that the selection of initial parameters of the DBN needs to rely on experience, Zheng et al. [92] provided a node recursion formula for the hidden layer in the DBN, which provided technical support for the selection of the number of hidden layer nodes, and constructed an aviation sensor fault diagnosis model based on the DBN, achieving good prediction results. Although DBN have absolute advantages in complex pattern recognition, fault classification, and feature extraction, the training process is complex and requires a lot of computing resources. In addition, the network structure and hyperparameter selection require many experiments for verification.

In summary, deep learning methods show significant advantages in automatically extracting complex features and processing large-scale data, as well as high diagnostic accuracy. However, these methods involve complex calculations during the training process and consume a large amount of computing resources. Therefore, they are particularly suitable for processing large-scale and diverse fault data, especially when the features are complex, and the patterns are nonlinear. The advantages, disadvantages, and usage scenarios of each fault diagnosis method based on deep learning are comparatively analyzed, as shown in Table 5. Future research will focus on optimizing algorithms to improve training efficiency and model robustness, solve small data problems, and explore the integration of deep learning and traditional signal processing technology to enhance the comprehensiveness and accuracy of fault diagnosis.

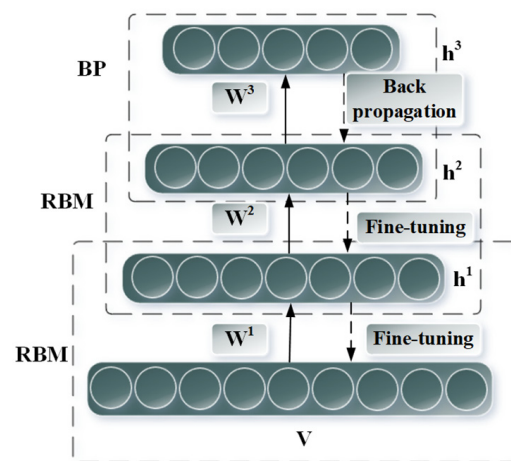


Figure 8. Schematic diagram of the deep belief network structure.

Table 5. Comparison of the advantages and disadvantages of various fault diagnosis methods based on deep learning.

Method	Advantage	Disadvantages	Applicable Scenarios
CNN	Good image and two-dimensional data processing effect, strong feature extraction ability	The input data preprocessing and normalization requirements are high, and the training resource requirements are large	Image data, time-frequency images, two-dimensional signal data
DBN	Automatically learn hierarchical feature representation, suitable for feature extraction of complex data	The training is complex, requires layer-by-layer training, and is sensitive to the choice of hyperparameters	Feature learning, data preprocessing
Recursive neural network	Processing of sequence data and time series analysis, capturing temporal dependencies	Training is unstable, and gradient disappearance and explosion problems are serious	Time series data, dynamic data analysis
Long short-term memory network	Capture long-term dependencies and perform well in processing long-sequence data	High computational complexity and long training time	Long-sequence time data analysis and fault prediction
Autoencoder	Feature dimensionality reduction, data compression, anomaly detection, unsupervised learning	There are many model structures and parameter choices, and the training process requires a lot of data	Feature learning, anomaly detection, data dimensionality reduction
Generative adversarial network	Generate high-quality samples suitable for data augmentation and anomaly detection	The training is complex and unstable, and the quality of the generated samples needs to be fine-tuned	Data enhancement, abnormal sample generation, generative model
Graph neural network	Processing structured graph data, considering node relationships and graph structure information	Graph structure construction and node feature selection require high computational complexity	Structured data, relationship network analysis

4.3. Comparative Analysis of Shallow Learning and Deep Learning

Based on the analysis of shallow learning and deep learning methods in Sections 4.1 and 4.2, their applications in fault diagnosis have their own advantages and limitations. A comparative analysis of shallow learning and deep learning is shown in Table 6. Shallow learning models (such as support vector machines, decision trees, and K-nearest neighbors) are usually suitable for scenarios with small data volumes, clear features, and high interpretability. Due to the low model complexity of shallow learning, its training and reasoning process is more efficient and relies on manually designed feature extraction. Through experts using prior knowledge, such as time-domain or frequency-domain analysis, to provide input features for the model, shallow learning can achieve good results when features are simple and prior knowledge is rich, such

as applying statistical features in the time-domain or frequency-domain in gear fault detection. However, such models have limited performance in dealing with high-dimensional data and complex nonlinear problems, and due to their reliance on manual feature extraction, feature quality has a decisive influence on model performance.

Table 6. Comparative analysis of shallow learning and deep learning.

Comparison Dimensions		Shallow Learning	Deep Learning
Model complexity and feature extraction	Advantages:	The model is simple, and the training and inference time is short.	The model is complex, and features are learned automatically; It has strong expressive power for complex data.
	Limitations:	Reliance on manual feature extraction	Training relies on a large amount of computing resources and data.
Data requirements	Advantages:	Suitable for scenarios with less data; Does not rely on a large amount of labeled data.	Excels in data-rich situations.
	Limitations:	Difficulty handling complex, nonlinear data.	Large-scale data is required.
Explanation and generalization	Advantages:	High transparency and easy to interpret.	Strong generalization ability.
	Limitations:	The generalization ability is weak, and it is difficult to handle complex nonlinear problems.	Poor interpretability, easily regarded as a “black box” model.
Training cost	Advantages:	The training and inference costs are low; Suitable for scenarios with limited resources and high real-time requirements.	It has powerful representation learning capabilities.
	Limitations:	For high-dimensional data, training and parameter adjustment are more complicated.	The training cost is high and the inference speed is slow; It requires a lot of computing resources.
Applicable scenarios		Small datasets, tasks with clear features; Industrial applications that require strong interpretability.	Large data sets, complex pattern recognition tasks; Automated fault detection with sufficient computing resources.

In contrast, deep learning performs well in large-scale data sets and complex pattern recognition and is particularly good at automatically extracting complex features from raw data. It is suitable for scenarios with abundant data volumes and difficult feature extraction, such as gear fault classification based on time-frequency diagrams of vibration signals. Deep learning models (such as convolutional neural networks and recurrent neural networks) have unique advantages for processing complex structured data. However, their complexity results in high computing resource requirements and training costs, and they are prone to overfitting problems when the amount of data is insufficient. In addition, the “black box” nature of deep learning reduces the interpretability of the model and may be limited in some industrial applications that require high interpretability. In general, the choice between shallow learning and deep learning should be weighed according to the specific application scenario and data size. The two can also complement each other through hybrid models to improve the performance of fault diagnosis.

5. Conclusions

Through the above analysis of the status of research, although the research on gear transmission system fault diagnosis methods has achieved many results, there are still many areas that need to be further explored. Improving and combining mathematical signal processing methods is still an important direction that needs in-depth research. At

the same time, combining mathematical signal processing methods with deep learning technology is also a key issue for future research. Future research directions and focuses include the following:

(1) Signal decomposition sparsity problem: The sparsity problem of the vibration signal decomposition results is particularly prominent when processing time-varying non-stationary signals. Existing decomposition methods often result in the number of signal components being significantly greater than the number of actual feature components of concern; that is, the decomposition results are often not sparse, and these components even contain false components that have no correlation with the real signal features, resulting in important fault characteristics being dispersed into multiple signal components, making it impossible to detect mechanical faults effectively and in a timely manner. However, existing fault diagnosis methods need to define appropriate indicators, select signal components related to the fault characteristics from the decomposition results, and appropriately combine these components to accurately reflect the mechanical fault characteristics. Therefore, suppressing the generation of false components, avoiding tedious component selection and combination steps, directly extracting the characteristic components of interest from the signal, and obtaining the sparsest signal decomposition results are some of the key issues for improving the performance of mechanical fault diagnosis.

(2) Separation of adjacent characteristic components: Mechanical vibration signals are highly complex, and some important characteristic components often have very close frequency intervals, such as the gear meshing frequency and its sideband components. Accurately identifying and separating these adjacent components is crucial for mechanical fault diagnosis. However, due to the inherent limitations of the algorithm, existing signal decomposition methods make it difficult to effectively process components with very close frequency distances. For example, when the frequency interval of signal components is small, the spectrum and time-frequency transform coefficients of different components interfere with each other, resulting in the inability of frequency-domain analysis methods and time-frequency-domain analysis methods to accurately separate these signal components. In addition, decomposition methods such as time-frequency analysis (EMD), which rely on signal extreme point detection, have performance that is heavily dependent on the amplitude and frequency ratios of the components. Breaking through the decomposition mode of existing methods and improving the resolution of signal decomposition methods are key issues for improving the performance of mechanical fault diagnosis.

(3) Problem of extracting weak signals: In the early stage of mechanical equipment failure, since the fault characteristic components are relatively weak, these characteristics in the vibration signal are easily submerged by noise and other irrelevant signal components. Therefore, accurately identifying and extracting weak characteristic components in vibration signals is crucial for the early fault diagnosis of mechanical equipment. However, when extracting weak signal components, signal decomposition methods are easily interfered with by other strong signal components, making it difficult to obtain satisfactory analysis results. For example, under strong interference, model optimization algorithms are prone to stability problems, resulting in divergent decomposition results; the analysis results of time-frequency-domain methods are related to the signal amplitude, and strong signal components will suppress the time-frequency representation results of weak signal components, making it difficult to identify weak signal components. Improving the robustness of signal decomposition methods and reducing the influence of signal amplitude on decomposition results is one of the key issues in improving the performance of signal decomposition methods for extracting weak signal components and mechanical fault diagnosis.

(4) Diagnosis and identification problems of complex faults: Complex fault diagnosis methods based on vibration signal processing usually focus on improving a specific signal decomposition technology, such as wavelet transform and empirical mode decomposition. However, the complexity of the components of complex fault signals and the influence of multi-source external excitations, such as random strong impact, dynamic load, and strong noise, make the diagnosis effect of traditional signal decomposition methods less

than ideal in practical applications. In addition, most intelligent complex fault diagnosis models are usually trained using only single fault data, which means that if the difference characteristics of two types of faults in a complex fault are not obvious, traditional machine learning models may find it difficult to accurately identify and distinguish these faults. Combining and optimizing traditional mathematical signal processing methods to develop a complex fault diagnosis model that can process complex signals and use multi-source data for training is one of the key issues in improving the performance of mechanical fault diagnosis.

(5) Multi-combination problem of fault diagnosis: To cope with complex and changeable fault signals, by combining artificial intelligence and machine learning technology, shallow learning and deep learning methods are integrated, and combined with optimization algorithms, an intelligent fault diagnosis system is constructed. The system is capable of self-learning and self-adaptation, and a special deep learning model is designed according to the characteristics of the vibration signals to realize the automatic extraction and identification of fault features. In addition, the small sample size problem is one of the current research focuses. The optimization algorithm can achieve fast and effective fault identification in the case of small sample data and improve the practicality and flexibility of the system. The effective combination of intelligent diagnosis technology with mathematical signal processing methods to improve the ability to identify fault signal features is one of the key issues in improving the performance of mechanical fault diagnosis.

(6) Multi-source data fusion technology: To achieve a more comprehensive and multi-level fault diagnosis, multi-channel sensors, such as vibration, temperature, acoustic emission, and image sensors can be combined to provide more comprehensive data. By using sensor fusion algorithms to process and integrate different types of data and give full play to the advantages of various sensors, the fusion of multi-source data can not only capture details that may be missed by a single sensor, but also improve the accuracy and reliability of fault diagnosis by comprehensively analyzing the associations between different data sources. In addition, multi-channel sensor fusion technology can also improve the robustness and anti-interference ability of the fault diagnosis system, making it adaptable to complex industrial environments.

Author Contributions: Article conception, X.F. and Y.F.; literature search, H.X. and N.P.; drafted and revised the work, X.F., Y.F., Y.X. and G.M. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Guangxi Natural Science Foundation (No. 2022GXNSFBA035574), the Guangxi Science and Technology Program (AD23026183), and the National Natural Science Foundation of China (No. 52265006).

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: Authors Xuezhong Fu, Guo Ma and Nanjiang Peng were employed by the Liuzhou Wuling Automobile Industry Co., Ltd. Author Haijun Xu was employed by the Liuzhou Wuling New Energy Automobile Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Ding, K.; Li, W.H.; Zhu, X.Y. *Practical Technology of Gear and Gearbox Fault Diagnosis*; Machinery Industry Press: Beijing, China, 2005; pp. 156–179.
2. Litvin, L.F.; Fuentes, A.; Gonzalez-Perez, I.; Carvenali, L.; Kawasaki, K.; Handschuh, R.F. Modified involute helical gears: Computerized design, simulation of meshing and stress analysis. *Comput. Methods Appl. Mech. Eng.* **2003**, *192*, 3619–3655. [[CrossRef](#)]
3. Kolivand, M.; Kahraman, A. An Ease-Off Based Method for Loaded Tooth Contact Analysis of Hypoid Gears Having Local and Global Surface Deviations. *J. Mech. Des.* **2010**, *132*, 071004. [[CrossRef](#)]
4. Fuentes-Aznar, A.; Gonzalez-Perez, I. Mathematical definition and computerized modeling of spherical involute and octoidal bevel gears generated by crown gear. *Mech. Mach. Theory* **2016**, *106*, 94–114. [[CrossRef](#)]

5. Maksat, T.; Christos, S.; Dongming, W. A computationally robust solution to the contact problem of two rotating gear surfaces in space. *Meccanica* **2023**, *58*, 2455–2466.
6. Hamza, T.; Zhaksylyk, G.; Andas, A.; Christos, S. Assessment of contact forces and stresses, torque ripple and efficiency of a cycloidal gear drive and its involute kinematical equivalent. *Mech. Based Des. Struct. Mach.* **2024**, *52*, 1304–1323.
7. Chen, P.; Toyota, T.; He, Z. Automated Function Generation of Symptom Parameters and Application to Fault Diagnosis of Machinery Under Variable Operating Conditions. *IEEE Trans. Syst. Man Cybern. Syst.* **2001**, *31*, 775–781. [[CrossRef](#)]
8. Chen, P.; Taniguchi, M.; Toyota, T.; He, Z. Fault diagnosis method for machinery in unsteady operating condition by instantaneous power spectrum and genetic programming. *Mech. Syst. Signal Process.* **2003**, *19*, 175–194. [[CrossRef](#)]
9. Zhong, B.L.; Huang, R. *Mechanical Fault Diagnosis*, 3rd ed.; Machinery Industry Press: Beijing, China, 2007; pp. 141–161.
10. Hua, W. Research on Vibration Fault Diagnosis of Low-Speed Heavy-Loaded Gears. Master's Thesis, Henan Polytechnic University, Nanyang, China, 2007.
11. Randall, R.B. New Method of Modeling Gear Faults. *ASME J. Mech. Des.* **1981**, *104*, 259–267. [[CrossRef](#)]
12. McFadden, P.D.; Wang, W.J. Early detection of gear failure by vibration analysis—II. Interpretation of the time-frequency distribution using image processing techniques. *Mech. Syst. Signal Process.* **1993**, *7*, 205–215.
13. Forrester, B.D. Gear Fault Detection using the Wigner-Ville Distribution. In *Australian Vibration and Noise Conference 1990, Vibration and Noise-Measurement Prediction and Control*; Preprints of Papers; Institution of Engineers: Barton, Australia, 1990.
14. Yu, H.P.; Dong, W.F.; Liu, Y. Application of CAT technology in gearbox noise testing. *China Test.* **2009**, *35*, 90–93.
15. Samanta, B. Gear fault detection using artificial neural networks and support vector machines with genetic algorithms. *Mech. Syst. Signal Process.* **2004**, *18*, 625–644. [[CrossRef](#)]
16. Cheng, G.; Cheng, Y.; Shen, L.; Qiu, J.; Zhang, S. Gear fault identification based on Hilbert–Huang transform and SOM neural network. *Measurement* **2013**, *46*, 1137–1146. [[CrossRef](#)]
17. Sanz, J.; Perera, R.; Huerta, C. Gear dynamics monitoring using discrete wavelet transformation and multi-layer perceptron neural networks. *Appl. Soft Comput. J.* **2012**, *12*, 2867–2878. [[CrossRef](#)]
18. Hu, P.; Zhao, C.; Huang, J.; Song, T. Intelligent and Small Samples Gear Fault Detection Based on Wavelet Analysis and Improved CNN. *Processes* **2023**, *11*, 2969. [[CrossRef](#)]
19. Bordoloi, D.; Tiwari, R. Optimization of support vector machine based multi-fault classification with evolutionary algorithms from time domain vibration data of gears. *Proc. Inst. Mech. Eng. Part C J. Mech. Eng. Sci.* **2013**, *227*, 2428–2439. [[CrossRef](#)]
20. Bond, R.R. *Vibration-Based Condition Monitoring: Industrial, Aerospace and Automotive Applications*; John Wiley & Sons, Ltd.: New York, NY, USA, 2010; pp. 143–165.
21. Simani, S.; Fantuzzi, C.; Patton, R.J. *Model-Based Fault Diagnosis in Dynamic Systems Using Identification Techniques*; Springer: Berlin/Heidelberg, Germany, 2003; pp. 61–113.
22. Fu, Y.; Chen, X.; Liu, Y.; Son, C.; Yang, Y. Multi-Source Information Fusion Fault Diagnosis for Gearboxes Based on SDP and VGG. *Appl. Sci.* **2022**, *12*, 6323. [[CrossRef](#)]
23. Zhang, X.; Wu, T.; Xu, J.; Li, B. Keyless phase time domain synchronous averaging method for variable speed gearbox vibration signal monitoring. *J. Xi'an Jiaotong Univ.* **2012**, *46*, 111–114.
24. Ha, M.J.; Youn, D.B.; Oh, H.; Han, B.; Jung, Y.; Park, J. Autocorrelation-based time synchronous averaging for condition monitoring of planetary gearboxes in wind turbines. *Mech. Syst. Signal Process.* **2016**, *70–71*, 161–175. [[CrossRef](#)]
25. Yao, G.; Wang, Y.; Benbouzid, M.; Ait-Ahmed, M. A Hybrid Gearbox Fault Diagnosis Method Based on GWO-VMD and DE-KELM. *Appl. Sci.* **2021**, *11*, 4996. [[CrossRef](#)]
26. Zhang, X.; Zhang, H. Application of Kurtosis-Based Deconvolution for Gear Fault Diagnosis. *Mech. Syst. Signal Process.* **2007**, *22*, 13–25.
27. Liu, H.; Wang, X. Deconvolution Approach Using Maximum Correlation Kurtosis for Gearbox Fault Diagnosis. *Mechanical Syst. Signal Process.* **2009**, *32*, 13–15.
28. Wu, Q.; Zhang, X. Enhanced Maximum Correlation Kurtosis Deconvolution Method for Gear Fault Detection. *IEEE Trans. Instrum. Meas.* **2011**, *9*, 73–95.
29. Hua, W.; Luo, C.; Leng, J.; Wang, Z. Mine gearbox fault diagnosis based on multiwavelets and maximum correlated kurtosis deconvolution. *J. Vibroeng.* **2017**, *19*, 4185–4197. [[CrossRef](#)]
30. Chen, L.; Xu, Y. Multi-Scale Maximum Correlation Kurtosis Deconvolution for Fault Diagnosis of Gearbox. *Mech. Syst. Signal Process.* **2017**, *42*, 43–65.
31. He, Q.; Guo, Y.G.; Wang, X.; Xie, P.; Jiang, G.; Li, J. Gearbox fault diagnosis based on signal resonance sparse decomposition and maximum correlation kurtosis deconvolution. *China Mech. Eng.* **2017**, *28*, 1528–1534.
32. Hong, L.; Liu, X.; Zuo, H. Compound Fault Diagnosis of Rotating Machinery Based on Adaptive Maximum Correlated Kurtosis Deconvolution and Customized Multiwavelet Transform. *Measurement* **2019**, *146*, 87–100. [[CrossRef](#)]
33. He, F.; Zheng, C.; Pang, C.; Zhao, C.; Yang, M.; Zhu, Y.; Luo, Z.; Luo, H.; Li, L.; Jiang, H. An Adaptive Deconvolution Method with Improve Enhanced Envelope Spectrum and Its Application for Bearing Fault Feature Extraction. *Sensors* **2024**, *24*, 951. [[CrossRef](#)]
34. Wang, K.X.; Liu, R.C.; Sun, J.Z. Gear fault monitoring method based on electrostatic signal and short-time Fourier transform. *Mech. Electr. Eng.* **2023**, *40*, 1664–1672.
35. Wu, H.; Lv, Y.; Yi, C.; Yuan, R. Synchronous compression algorithm with window expansion optimization and its application in instantaneous frequency estimation under variable speed conditions. *China Mech. Eng.* **2022**, *33*, 34–44.

36. Bao, W.; Tu, X.; Li, F.; Dai, L.; Shen, Y.; Liu, Z.; Wang, X. Parameterized short-time Fourier transform and gearbox fault diagnosis. *Vibration. Test Diagn.* **2020**, *40*, 272–277.
37. Zhang, Z.L. Research on Feature Detection of Time-Varying Non-Stationary Signals Based on Fractional Order Transform. Master's Thesis, Beijing Jiaotong University, Beijing, China, 2021.
38. Jiao, L. Research on Non-Stationary Early Fault Diagnosis of Planetary Gears. Master's Thesis, Beijing Information Science and Technology University, Beijing, China, 2021.
39. Hu, R.; Pang, X.; She, C.; Liu, A.; Hu, S. Fault diagnosis of planetary gearbox under variable working conditions based on optimal window function Gabor transform. *Wind Turbine Technol.* **2021**, *63*, 79–90.
40. Zhao, H.; Zhang, Z.; Mei, J.; Shen, H.; Chang, C. Early fault diagnosis of transmission gear based on FRFT and LSTM. *J. Mil. Transp. Acad.* **2020**, *22*, 36–41.
41. Wang, H. Fractional S transform and its application in gearbox fault diagnosis. *J. Electron. Meas. Instrum.* **2019**, *33*, 133–139.
42. Zhang, Z.; Mei, J.; Chang, C.; Zhao, H.; Shen, H. Feature extraction of early fault of gear based on FRFT target separation and MCKD feature enhancement. *J. Mil. Transp. Acad.* **2019**, *21*, 44–48.
43. Guo, M.; Zhou, C.; Zheng, C.; Chen, P.; Hu, G. Variable working condition gear fault diagnosis method based on meta-learning technology. *Mech. Electr. Eng.* **2023**, *40*, 1682–1690.
44. Yu, C.; Liang, R.; Ran, W.; Wang, Z. Gearbox fault diagnosis based on STFT and CNN. *Mech. Manuf. Autom.* **2022**, *51*, 152–154.
45. Chen, X.; Liu, Q.J.; Wang, Z.Y. Research on transmission system fault diagnosis based on FFT and LLTSA. *Mech. Electr. Eng.* **2022**, *39*, 513–518.
46. Zhen, C.; Tian, Y.; Ji, K.; Zhang, Z.; Huang, D. Gearbox oil temperature data prediction based on FFT and DNN. *Comput. Mod.* **2022**, *4*, 17.
47. Zhu, Q.; He, H.; Li, F.; Li, Z.; Li, Z. Helicopter gearbox fault diagnosis method based on short-time Fourier transform and deep convolutional neural network. *Fail. Anal. Prev.* **2022**, *17*, 1–8.
48. Zhang, L.; Xu, W.; Jing, L.; Tan, J. Rotating machinery fault diagnosis based on two-dimensional deep convolutional network. *Mech. Strength* **2020**, *42*, 1039–1044.
49. Morlet, J.; Arens, G.; Fourgeau, E.; Glard, D. Wave Propagation and Sampling Theory-Part I: Complex Signal and Scattering in Multilayered Media. *Geophysics* **2013**, *47*, 203–221. [[CrossRef](#)]
50. Mallat, S.G. A Theory for Multiresolution Signal Decomposition: The Wavelet Representation. *IEEE Trans. Pattern Anal. Mach. Intell.* **1989**, *11*, 674–693. [[CrossRef](#)]
51. Donoho, D.L.; Johnstone, I.M. Ideal Spatial Adaptation by Wavelet Shrinkage. *Biometrika* **1994**, *81*, 425–455. [[CrossRef](#)]
52. Parey, A.; Singh, A. Gearbox Fault Diagnosis Using Acoustic Signals, Continuous Wavelet Transform and Adaptive Neuro-Fuzzy Inference System. *Appl. Acoust.* **2019**, *147*, 133–140. [[CrossRef](#)]
53. Liang, P.; Deng, C.; Wu, J.; Yang, Z.; Zhu, J.; Zhang, Z. Compound Fault Diagnosis of Gearboxes Via Multi-label Convolutional Neural Network and Wavelet Transform. *Comput. Ind.* **2019**, *113*, 103–132. [[CrossRef](#)]
54. Saravanan, N.; Ramachandran, I. Fault Diagnosis of Spur Bevel Gear Box Using Discrete Wavelet Features and Decision Tree Classification. *Expert Syst. Appl.* **2009**, *36*, 9564–9573. [[CrossRef](#)]
55. Huang, N.E.; Shen, Z.; Long, S.R.; Wu, M.C.; Shih, H.; Zheng, Q.; Yen, N.C.; Tung, C.C.; Liu, H.H. The Empirical Mode Decomposition and the Hilbert Spectrum for Nonlinear and Nonstationary Time Series Analysis. *Proc. R. Soc. A* **1998**, *454*, 903–995. [[CrossRef](#)]
56. Wu, Z.; Huang, N. Ensemble Empirical Mode Decomposition: A Noise-Assisted Data Analysis Method. *Adv. Adapt. Data Anal.* **2009**, *1*, 1–41. [[CrossRef](#)]
57. Zhang, C.; Yuan, Y.X. Application of frequency modulation empirical mode decomposition in bearing fault diagnosis. *J. Vib. Shock* **2014**, *33*, 185–189.
58. Park, S.; Kim, S.; Choi, J. Gear fault diagnosis using transmission error and ensemble empirical mode decomposition. *Mech. Syst. Signal Process.* **2018**, *108*, 262–275. [[CrossRef](#)]
59. Tang, B.P.; Jiang, Y.H.; Zhang, X.C. Rolling bearing fault feature extraction method based on morphological singular value decomposition and empirical mode decomposition. *J. Mech. Eng.* **2010**, *46*, 37–42. [[CrossRef](#)]
60. Yang, Q.; Yang, Q.; Huang, M.; Huang, M.; Yan, W. Particle swarm optimization-based empirical mode decomposition–kernel independent component analysis joint approach for diagnosing wind turbine gearbox with multiple faults. *Trans. Inst. Meas. Control* **2018**, *40*, 1836–1845. [[CrossRef](#)]
61. Zhang, L.W.; Qi, F.; Lin, J.; Zhang, Z.C.; Yan, X.H.; Li, C. A nearly end-to-end deep learning approach to fault diagnosis of wind turbine gearboxes under nonstationary conditions. *Eng. Appl. Artif. Intell.* **2023**, *119*, 105735. [[CrossRef](#)]
62. Li, Y.; Si, S.; Liu, Z.; Liang, X. Review of local mean decomposition and its application in fault diagnosis of rotating machinery. *J. Syst. Eng. Electron.* **2019**, *30*, 799–814.
63. Niu, X.; Lu, L.; Wang, J.; Han, X.; Li, X. An improved empirical mode decomposition based on local integral mean and its application in signal processing. *Math. Probl. Eng.* **2021**, *2021*, 8891217. [[CrossRef](#)]
64. Cheng, J.S.; Zheng, J.D.; Yang, Y. A new method for non-stationary signal analysis: Local characteristic scale decomposition method. *J. Vib. Eng.* **2012**, *25*, 215–220.
65. Cheng, J.; Yang, Y.; Yang, Y. A rotating machinery fault diagnosis method based on local mean decomposition. *Digit. Signal Process.* **2012**, *22*, 356–366. [[CrossRef](#)]

66. Pan, Y.; Hu, Z.; Chen, J.; Wang, H.; Hong, R. Improved Complete Ensemble Robust Local Mean Decomposition with Adaptive Noise for Slewing Bearings Performance Degradation Assessment. *IEEE Access* **2022**, *10*, 78677–78690. [\[CrossRef\]](#)
67. Zhou, J.; Cheng, J.; Wu, X.; Wang, J.; Cheng, J.; Yang, Y. Completely adaptive projection multivariate local characteristic-scale decomposition and its application to gear fault diagnosis. *Measurement* **2022**, *202*, 111743. [\[CrossRef\]](#)
68. Cui, W.; Xu, A.; Li, W.; Li, W.; Meng, F. Gear fault diagnosis based on local characteristic scale decomposition combined with local mean demodulation. *China Mech. Eng.* **2016**, *27*, 3332–3337.
69. Goyal, D.; Choudhary, A.; Sandhu, J.K.; Srivastava, P.; Saxena, K.K. An intelligent self-adaptive bearing fault diagnosis approach based on improved local mean decomposition. *Int. J. Interact. Des. Manuf.* **2022**, 1–11. [\[CrossRef\]](#)
70. Narendra, K.S.; Parthasarathy, K. Identification and control of dynamical systems using neural networks. *IEEE Trans. Neural Netw.* **1990**, *1*, 4–27. [\[CrossRef\]](#) [\[PubMed\]](#)
71. Yu, H.J. Artificial neural network identification method for mechanical equipment fault diagnosis. *Mech. Strength* **1995**, *17*, 48–54.
72. Liao, X.X.; Chang, L.; Shen, Y. Research on the stability of discrete Hopfield neural network. *Chin. J. Autom.* **1999**, *25*, 721–727.
73. Jin, S.; Cui, W.; Jin, Z.; Wang, Y. AF-DHNN: Fuzzy Clustering and Inference-Based Node Fault Diagnosis Method for Fire Detection. *Sensors* **2015**, *15*, 17366–17396. [\[CrossRef\]](#)
74. Skowron, M.; Wolkiewicz, M.; Orłowska, K.T.; Kowalski, C.T. Effectiveness of Selected Neural Network Structures Based on Axial Flux Analysis in Stator and Rotor Winding Incipient Fault Detection of Inverter-fed Induction Motors. *Energies* **2019**, *12*, 2392. [\[CrossRef\]](#)
75. Lang, S.C. Research on Ship Power Grid Fault Diagnosis Based on Hopfield Neural Network. Master's Thesis, Chongqing Jiaotong University, Chongqing, China, 2019.
76. Lin, T.; Yang, X.; Cai, R.; Zhang, L.; Liu, G. Elman neural network wind turbine fault diagnosis based on improved artificial bee colony algorithm. *Renew. Energy* **2019**, *37*, 612–617.
77. Aydmj, T.; Duin, R. Pump failure determination using support vector data descriptions. *Lect. Notes Comput. Sci.* **1999**, *8*, 415–425.
78. Tang, B.; Song, T.; Li, F.; Deng, L. Fault diagnosis for a wind turbine transmission system based on manifold learning and Shannon wavelet support vector machine. *Renew. Energy* **2014**, *62*, 1–9. [\[CrossRef\]](#)
79. Yang, D.; Liu, Y.; Li, S.; Li, X.; Ma, L. Gear fault diagnosis based on support vector machine optimized by artificial bee colony algorithm. *Mech. Mach. Theory* **2015**, *90*, 219–229. [\[CrossRef\]](#)
80. Martin, D.I.; Morinigo, S.D.; Duque, P.O.; Romero-Troncoso, R.D.J. Early fault detection in induction motors using AdaBoost with imbalanced small data and optimized sampling. *IEEE Trans. Ind. Appl.* **2016**, *53*, 3066–3075. [\[CrossRef\]](#)
81. Ren, Z.; Huang, H.; Hu, Y. Research on gear fault diagnosis based on improved support vector machine. *Mech. Des. Manuf.* **2022**, *4*, 122–125.
82. Hubel, D.H.; Wiesel, T.N. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Physiology* **1962**, *160*, 106–154. [\[CrossRef\]](#)
83. Dong, Y.; Wang, L.; Zhang, C.; Zhao, Q. Gearbox health status identification based on deep convolutional neural network. *Comput. Simul.* **2024**, *41*, 455–459.
84. Zhang, Y.; Zhao, X.; Hui, Y.; Chen, P. SA-DACNN gearbox fault diagnosis method based on multi-sensor data fusion. *Control Decis.* **2024**, *45*, 60–71.
85. Huang, K.; Wu, B.; Zhang, Z. Intelligent diagnosis method of planetary gear based on integrated convolutional neural network. *Mech. Des. Manuf.* **2024**, *1*, 170–174.
86. Zhou, Q.; Li, Y.; Tian, Y.; Jiang, L. A novel method based on nonlinear auto-regression neural network and convolutional neural network for imbalanced fault diagnosis of rotating machinery. *Measurement* **2020**, *161*, 107880. [\[CrossRef\]](#)
87. Liang, X.; Zuo, J.; Hoseini, R. Vibration Signal Modeling of a Planetary Gear Set for Tooth Crack Detection. *Eng. Fail. Anal.* **2015**, *48*, 185–200. [\[CrossRef\]](#)
88. Tamilselvan, P.; Wang, P.F. Failure diagnosis using deep belief learning based health state classification. *Reliab. Eng. Syst. Saf.* **2013**, *115*, 124–135. [\[CrossRef\]](#)
89. Yin, J.T.; Zhao, W.T. Fault diagnosis network design for vehicle on-board equipment of high-speed railway: A deep learning approach. *Eng. Appl. Artif. Intell.* **2016**, *56*, 250–259. [\[CrossRef\]](#)
90. Zhang, C.; He, Y.; Jiang, S.; Wang, T.; Yuan, L.; Li, B. Transformer fault diagnosis method based on self-powered RFID sensor tag, DBN, and MKSVM. *IEEE Sens. J.* **2019**, *19*, 8202–8214. [\[CrossRef\]](#)
91. Yan, L.; Dong, X.; Zhang, Y.; Wang, T.; Gao, Q. Gas turbine gas path fault diagnosis method based on deep belief network. *J. Eng. Thermophys.* **2020**, *41*, 840–844.
92. Zheng, X.; Guo, C.; Yao, B.; Feng, H. Aviation sensor fault diagnosis method based on deep learning. *Comput. Eng.* **2017**, *43*, 281–287.

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