Proceedings of the 4th International Conference on Machine Learning and Automation
DOI: 10.54254/2755-2721/2026.37484

Deep Learning-Based Methods for Rolling Bearing Fault Diagnosis: Optimization and Comparison

Zexin Yin

Beijing 21st Century School, Beijing, China
allen35531@outlook.com

Abstract. Rolling bearings are the vital support elements of industrial rotating machinery. Operational conditions have a direct impact on the safety, dependability and economical operation and maintenance of the production system. An effective and accurate fault diagnosis technology is very important for predictive maintenance and reducing the probability of downtime. With the development of deep learning and signal processing technologies, intelligent fault diagnosis based on vibration signals has become a research hotspot, and various technical methods and implementation schemes have been proposed. In this paper, three traditional deep-learning-based approaches for the diagnosis of rolling bearing faults are systematically studied, including the time-frequency image diagnosis method based on the ICEEMDAN decomposition integrated with the Hilbert transform and CBAM - ResNet, the FSCL diagnosis model combining FFT + SVD with CNN - LSTM, and the classification framework fusing the continuous wavelet transform signal-to-image conversion with SE-ResNet. This paper also discusses the technical frameworks and effectiveness of each approach. In this study, the inherent limitations of the three approaches were systematically studied from pre-processing efficiency, model complexity, end-user adaptability and generalization ability of scenarios, and relevant optimization strategies were put forward. The research results can be used as reference to promote the technology progress and industrial application of intelligent fault diagnosis for rolling bearings.

Keywords: Rolling bearings, deep-learning-based approaches, optimization strategies

1. Introduction

Rolling bearings are the main supporting parts of industrial rotating equipment. Their operating status directly determines the reliability of the production system and the economic efficiency of operation and maintenance. The classic research by Randall et al. presented that bearing failures are the main cause of unplanned shutdowns of rotating machinery, accounting for more than 40% of all faults [1]. As industrial equipment develops towards higher speeds, heavier loads, and long-term continuous operation, the traditional maintenance model cannot balance cost control and risk prevention. Intelligent fault diagnosis based on vibration signals has become the core development direction in the field of predictive maintenance [2, 3].

The intelligent fault diagnosis of present rolling bearings is moving from laboratory verification to industrial applications. However, the limited edge computing power, insufficient generalization

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2026.37484

ability in different operating conditions, lack of fault samples and high real-time requirements have led the academic community to research in several directions: Li et al. proposed a lightweight transfer learning framework based on knowledge distillation to balance the diagnostic accuracy and deployment efficiency in cross-device scenarios [4]. Huang et al. proposed a causal decoupling domain adaptive model, which effectively enhanced the diagnostic robustness under different operating conditions [5]. Shao et al. built a deep learning framework incorporating reinforcement learning, which kept high accuracy even with small samples [6]. Kaya et al. proposed a quantized 1-D convolutional neural network for millisecond-level inference on microcontrollers, demonstrating the feasibility of real-time diagnosis at the edge [7].

Existing studies are mostly aimed at the performance improvement of a single method, and there is still a lack of systematic comparison on the inherent limitations, applicable boundaries and optimization paths of different technical routes. Therefore, this paper reviews three mainstream deep learning methods for rolling bearing fault diagnosis systematically. It carries out limitation analysis from the dimensions of signal preprocessing, model architecture, experimental verification, and generalization ability, and proposes corresponding optimization paths. It is hoped that it can provide reference support for intelligent rolling bearing diagnosis technology iteration and industrial implementation.

2. Related research on rolling bearing fault diagnosis cleaning robots

2.1. Case 1: method of time-frequency image diagnosis

Liang et al. proposed a novel approach for rolling bearing fault diagnosis based on ICEEMDAN and ResNet [8]. The method decomposes the bearing vibration signal by ICEEMDAN to obtain IMF components. Then, the one-dimensional time-domain signal is transformed into a two-dimensional time-frequency image through Hilbert transformation. Then, the fault features are extracted by a ResNet network with CBAM attention mechanism fusion. Transfer learning is used to solve the problem of insufficient fault samples in industrial scenarios. The method achieved an accuracy rate of 95.2% in the 10-class fault identification task of the CWRU bearing dataset, which is more than 12% higher than traditional models such as VGG and AlexNet, effectively improving the accuracy and robustness of fault.

2.2. Case 2: FSCL diagnostic model based on FFT-SVD and CNN-LSTM

Wu et al. proposed the FSCL fault diagnosis model, which combines traditional signal processing with deep learning [9]. It is helpful to improve the problems of strong noise, variable working conditions and unbalanced samples in industrial site bearing vibration signals. This model extracts the multi-dimensional fault features of FFT and SVD, and adopts CNN-LSTM network for local feature extraction and temporal dependency encoding to accurately identify fault types. The diagnostic accuracy rate of the model achieved 99.32% verified on the two public datasets of CWRU and XJTU-SY. It outperforms the mainstream algorithms, such as SVM and basic CNN, under complex conditions, including working condition switching, imbalanced samples, and strong noise interference. It has high accuracy and strong robustness, and can be used in the online and offline diagnosis of bearing fault.

DOI: 10.54254/2755-2721/2026.37484

2.3. Case 3: signal-to-image transformation and ResNet family based diagnostic scheme

Xu et al. proposed a rolling bearing fault diagnosis scheme based on signal to image conversion, which made full use of mature network architecture in the domain of computer vision [10]. In this paper, the one-dimensional vibration signals are transformed into two-dimensional time-frequency images using the continuous wavelet transform and the optimal wavelet basis function gaus1 is selected for comparison. The system tested five classic residual network models from ResNet18 to ResNet152 and introduced the channel attention mechanism to build the SE-ResNet series of models for control experiments. The CWRU bearing data set was used to verify for a number of single load and mixed load cases. The results showed that SE-ResNet152 had the best diagnostic performance, with an accuracy of 96.42%. The SE module has greatly improved the diagnostic effect of the model, which provides a systematic reference for the selection of deep learning models for bearing fault diagnosis.

Overall, the contents of three representative works are shown in Figure 1.

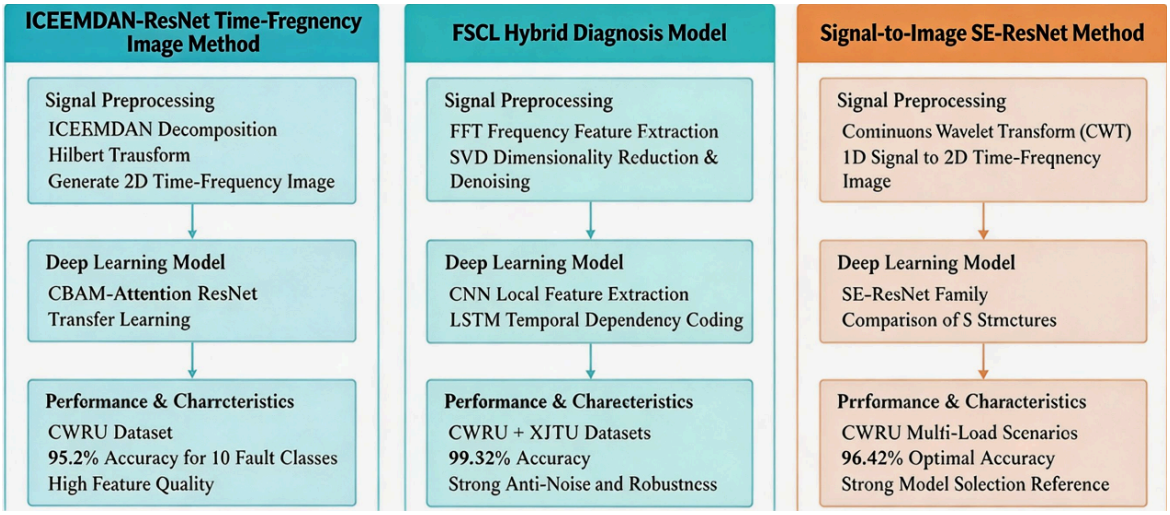


Figure 1. Core architecture of three representative deep learning fault diagnosis methods for rolling bearings (picture credit: original)

3. Analysis of limitations and solutions

3.1. Limitations and optimization strategies of signal preprocessing and feature extraction methods

Liang et al. adopted the ICEEMDAN decomposition with the Hilbert transform preprocessing scheme, which enhanced the quality of features. However, the decomposition process requires several times of adaptive noise superposition and iterative filtering, which leads to high computational complexity and long processing time. It is difficult to meet the requirements of real-time online diagnosis for industrial applications. For strong noise environment, the interference of noise is still present in the IMF components obtained by decomposition, and the false modes and mode aliasing cannot be completely removed. The Hilbert transform has the inherent endpoint effect, which distorts the edge features of the time-frequency images as well. Furthermore, the diagnostic logic is untraceable due to the "black box" of the deep residual network. transfer learning is the only way to solve the problem of the small sample size. Transfer learning can only solve the problem of small sample size, but the performance of transfer learning will degrade significantly

DOI: 10.54254/2755-2721/2026.37484

with large deviation among different operating conditions. Moreover, it has been only tested on artificial single-point fault datasets under laboratory conditions and has not been sufficiently validated on real industrial cases.

Regarding the mentioned problem of pre-processing efficiency, the iterative logic of ICEEMDAN can be improved by choosing suitable superimposition times of noise and threshold for component selection. Moreover, the decomposition speed can be enhanced by integrating with a parallel processing architecture. Before the decomposition, a pre-wavelet denoising method is added. The component correlation coefficient method is used to automatically select the effective IMF components after decomposition. The Hilbert endpoint effect is removed by using the signal extension algorithm. As for the model limitations and verification, the interpretability analysis methods can be introduced to clarify the basis for the diagnostic decisions. The transfer learning strategy can be improved to deal with larger domain shifts. Meanwhile, verification datasets are enriched by natural wear, compound faults and variable operating condition datasets.

3.2. Constraints and enhancement of the FSCL fault diagnosis framework

Wu et al.'s study used the FSCL model, which uses the feature extraction method based on the combination of FFT and SVD. It has inherent performance limitations based on its fundamental principles. FFT is a global frequency domain statistical analysis method, which can only reflect the overall frequency distribution of the signal, but cannot reflect the instantaneous time-varying characteristics of non-stationary vibration signals. It lacks a proper understanding of the effect attributes of the initial small faults in the bearing, and it is easy to cause premature misdiagnosis. The effectiveness of SVD dimensionality reduction is evaluated in an experiential way. The quality of features is greatly dependent on parameter optimization. Moreover, there is usually redundancy and collinearity of information in a large number of extracted features, which confuses the boundary of fault classification. Meanwhile, the CNN-LSTM hybrid model has a large number of parameters and high inference costs, which is difficult to deploy on industrial edge devices with limited computing power. The "black box" nature of the model makes it impossible to trace the diagnostic reasoning and causes a significant loss of precision when switching between operational states. The validation of real industrial composite fault scenarios is lacking.

To remedy the above shortcomings of preprocessing, short-time Fourier transform or continuous wavelet transform may be used to generate time-frequency joint features, which may enhance the local time-varying data and improve the detection of nascent weak faults. The energy ratio approach can be combined with mutual information feature selection for adaptive selection of singular value orders to remove redundant dimensions with large collinearity. For the practical application and deployment of models, lightweighting of networks can be achieved by model pruning and knowledge distillation for the implementation on the edge side. To alleviate the distribution shifts in working conditions, domain adaptation techniques can be applied and interpretability components can be added to keep track of diagnostic evidence. Simultaneously, for verification, the data of real industrial environments such as natural degradation and composite failures should be included, thus enhancing comprehensively. In the meantime, real industrial scenarios such as natural wear and compound faults should be added for validation, which comprehensively improves the robustness of the model in practical deployment.

DOI: 10.54254/2755-2721/2026.37484

3.3. Limitation and optimization routes for signal-to-image and residual network diagnostic schemes

Xu et al. proposed a scheme which transforms the one-dimension vibration signals into two-dimension time-frequency images by continuous wavelet transform. Although it reuses mature image classification networks, its computational complexity is much higher than that of traditional frequency-domain methods, and the processing of long-duration signals takes a long time, making it difficult to meet real-time diagnostic requirements in industrial environments. The study just fixedly selects the gaus1 wavelet basis function, and cannot be adaptively adjusted according to fault types and operating conditions. The transient impact characteristics of early weak faults are easily smoothed and weakened during the process of signal to image conversion, which reduces the detection rate of early faults. Meanwhile, the scheme only adopts the SE attention mechanism in the channel dimension, which cannot accurately focus on the spatial location of fault features. And it only picks models from the ResNet family, which leads to many optimal model parameters and slow inference speed. It is only validated on the single CWRU artificial fault dataset, and does not cover real-world scenarios such as compound faults and strong noise, and its generalization capability is not systematically verified.

To rectify the deficiencies in the aforementioned signal conversion, the implementation of fast wavelet transform alongside parallel computing may be utilized to improve processing efficacy. An adaptive selection system for wavelet bases can be developed, automatically aligning the ideal basis function with the signal's noise level and fault attributes. Prior to conversion, preprocessing for feature enhancement of the signal is conducted to emphasize the influence characteristics of nascent minor faults and minimize the loss of mapping data. To address the constraints of the network and its implementation, dual-dimensional attention modules for channels and spaces can be incorporated to accurately identify fault characteristics. Lightweight network architectures and model compression methods are presented to diminish the parameter count and inference expenses, rendering them appropriate for implementation on edge devices. At the same time, a multi-source real fault data set is supplemented for verification, and transfer learning and domain adaptation algorithms are combined to improve the generalization ability of small samples and cross-condition robustness

In summary of the above analysis and discussion, Table 1 lists the main limitations of the three research works and the corresponding suggestions, as detailed below.

Table 1. three major limitations of the study and corresponding recommendations

Method	Main limitations	Optimization Suggestions
ICEEMDAN-ResNet [8]	1. High computational complexity of decomposition, poor real-time performance 2. Mode mixing under strong noise; endpoint distortion from Hilbert transform 3. Weak model interpretability; only validated on artificial single-point faults	1. Optimize iteration logic; add pre-denoising and signal extension 2. Improve interpretability and transfer learning for domain shift 3. Supplement natural wear and composite fault datasets

DOI: 10.54254/2755-2721/2026.37484

Table 1. (continued)

FSCL Model [9]	1. FFT lacks time-varying features; weak perception of early faults 2. Manual SVD order selection; feature collinearity blurs classification boundaries 3. Heavy CNN-LSTM architecture; obvious accuracy drop across working conditions	1. Add time-frequency features; adaptively select SVD order to remove collinearity 2. Lighten the network via pruning; introduce domain adaptation 3. Validate with real industrial scenario data
Signal-to-Image SE-ResNet [10]	1. High CWT computation cost; fixed wavelet basis weakens early features 2. Only channel attention; large optimal model with slow inference 3. Single dataset validation; insufficient generalization verification	1. Adopt fast CWT; build adaptive wavelet basis selection 2. Add spatial attention; apply lightweight model compression 3. Validate with multi-source real datasets to improve generalization

4. Conclusion

This paper systematically reviews and compares three mainstream deep learning-based rolling bearing fault diagnosis methods, elaborates on the technical approaches and performances of each method, analyzes the inherent limitations of each method from four dimensions: signal preprocessing, model architecture, experimental validation and generalization ability, and puts forward targeted optimization suggestions.

The three approaches differ in technical focus and the contexts in which they are relevant: The ICEEMDAN-ResNet time-frequency image approach has the advantages of robust feature representation and excellent diagnosis accuracy. The FSCL hybrid model has the advantages of strong noise immunity, good adaptability to operation environments, and higher computational efficiency. The signal-to-image SE-ResNet approach has the advantage of a well-established architecture and can serve as a good reference for model selection. However, all three methods have similar problems: insufficient real-time efficiency, great difficulties in deployment on the end-user side and limited generalization capability in real industrial scenarios. To overcome these obstacles, we need to continue to improve the signal processing algorithms, promote the model lightweighting, implement the domain adaptive learning, and enhance the multi-scenario validation.

In the future, the research should be targeted to the actual needs of industrial sites, such as lightweight real-time diagnosis at the edge, few-shot cross-working-condition generalization, and multi-source sensor information fusion, so as to further improve the engineering application value of deep learning fault diagnosis methods.

References

[1] Randall, R. B., & Antoni, J. (2011). Rolling element bearing diagnostics-A tutorial. *Mechanical Systems and Signal Processing, 25*(2), 485–520.

[2] Zhang, X., Zhao, B., & Lin, Y. (2021). Machine learning based bearing fault diagnosis using the Case Western Reserve University data: A review. *IEEE Access, 9*, 155598–155608.

[3] Khan, S., & Yairi, T. (2018). A review on the application of deep learning in system health management. *Mechanical Systems and Signal Processing, 107*, 241–265.

DOI: 10.54254/2755-2721/2026.37484

[4] Li, Y., Wang, H., & Zhang, L. (2024). Lightweight knowledge distillation-based transfer learning framework for rolling bearing fault diagnosis. *Sensors, 24*(5), Art. no. 1758.
[5] Huang, A., Li, X., Sun, S., et al. (2025). Causal disentanglement-based domain adaptation for rolling bearing fault diagnosis. *Journal of Zhejiang University (Engineering Science), 59*(7), 1523–1531.
[6] Shao, H., Han, G., & Shi, H. (2026). An integrated deep learning model with reinforcement learning for rolling bearing fault diagnosis under small-sample conditions. *J. Vib. Acoust., 148*(2), Art. no. 021008.
[7] Kaya, M., Tekin, R., & Eren, L. (2026). Effective classification of bearing vibration signals using supervised machine learning for predictive maintenance: A lightweight 1D CNN for embedded deployment. *J. Electron. Sci. Technol. Appl., 5*(1), Art. no. 590423.
[8] Liang, B., & Feng, W. (2023). Bearing fault diagnosis based on ICEEMDAN deep learning network. *Processes, 11*(8), Art. no. 2440.
[9] Wu, G., Ji, X., Yang, G., et al. (2023). Signal-to-image: Rolling bearing fault diagnosis using ResNet family deep-learning models. *Processes, 11*(5), Art. no. 1527.
[10] Xu, M., Yu, Q., Chen, S., et al. (2024). Rolling bearing fault diagnosis based on CNN-LSTM with FFT and SVD. *Information, 15*(7), Art. no. 399.