

Article

A Bearing Fault Diagnosis Method Integrating the SWT and MCNN—RIME—KELM Hybrid Model

Liping Wang, Xing Liu, Xiaoke Su and Dongyao Zou * 

School of Computer Science and Artificial Intelligence, Zhengzhou University of Light Industry, Zhengzhou 450000, China; kinglp2026@163.com (L.W.); 332305090546@zzuli.edu.cn (X.L.); suxiaoke07@126.com (X.S.)

* Correspondence: zdy@zzuli.edu.cn

Abstract

To address the issues of severe noise interference, limited classification capability of linear classifiers, and difficulty in adaptively optimizing classifier parameters in rolling bearing fault diagnosis, this paper proposes a hybrid diagnostic model integrating the multi-scale convolutional neural network and rime ice optimization algorithm optimized kernel extreme learning machine. The method first employs the synchrosqueezed wavelet transform to convert raw vibration signals into high-resolution time-frequency images, effectively enhancing the visualization of fault impact features. Then, the multi-scale convolutional neural network is used to extract preliminary features from the time-frequency images, and the kernel extreme learning machine is introduced to replace the Softmax linear classifier in traditional convolutional neural networks, thereby constructing a nonlinear decision boundary to more effectively separate complex fault patterns. Finally, the rime algorithm is introduced to optimize the regularization coefficient and kernel parameters of the kernel extreme learning machine, enabling the kernel extreme learning machine to perform fault classification with an optimal nonlinear decision boundary. Experimental results on the bearing datasets from Huazhong University of Science and Technology and Case Western Reserve University show that the proposed method achieves classification accuracies of 99.75% and 99.83%, respectively, outperforming several comparison models. Furthermore, noise robustness experiments demonstrate that the proposed model maintains an accuracy of approximately 90% under low signal-to-noise ratio (SNR) conditions, outperforming all comparison models and demonstrating high classification accuracy under strong noise.

Keywords: rolling bearing; fault diagnosis; SWT; MCNN; RIME; KELM



Academic Editor: Xuefang Xu

Received: 19 May 2026

Revised: 15 June 2026

Accepted: 16 June 2026

Published: 18 June 2026

Copyright: © 2026 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 \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Rolling bearings are the most fault-prone critical components in rotating machinery, and their operating status directly determines equipment reliability and safety. According to statistics, approximately 30% of rotating machinery failures is caused by bearing issues. In critical components such as gearboxes, this proportion can reach as high as 76%. Traditional signal analysis methods for rolling bearing fault diagnosis are mainly based on time-domain, frequency-domain, and time-frequency domain features. The short-time Fourier transform (STFT) performs time-frequency joint analysis via windowed segmentation. However, because of the Heisenberg uncertainty principle, it cannot achieve optimal time and frequency resolution simultaneously. The continuous wavelet

transform (CWT) achieves multi-resolution analysis through scale-variable basis functions, but its effectiveness heavily depends on the selection of wavelet basis and threshold, and its energy concentration remains insufficient [1]. These limitations lead to energy diffusion and blurred ridges in time-frequency images generated by traditional methods, making it difficult to identify transient impact features caused by bearing faults clearly. Daubechies et al. [2] proposed the synchrosqueezed wavelet transform (SWT) within the CWT framework. This transform compresses wavelet coefficients along the frequency direction while preserving signal reconstructability, thereby greatly enhancing the concentration of time-frequency representation. Subsequent studies have further implemented synchrosqueezing transforms within the STFT framework and developed second-order and higher-order synchrosqueezing transforms to handle more complex time-varying signals. With its excellent energy concentration capability, SWT can effectively suppress noise and interference, highlight the time-frequency ridges caused by fault impacts, and is thus highly suitable for fault feature extraction from non-stationary vibration signals. The synchrosqueezing method has been widely used in mechanical fault diagnosis and other fields due to its powerful energy focusing advantage [3–5]. Although traditional fault diagnosis can achieve relatively high diagnostic accuracy under certain conditions, it relies on manual feature extraction, which has problems such as strong subjectivity and poor noise resistance [6].

Deep learning methods have been widely introduced into the field of rolling bearing fault diagnosis due to their powerful feature self-learning capabilities. Convolutional neural networks (CNN) can automatically extract discriminative features from time-frequency images, avoiding the tedious process of manual feature engineering and achieving better results than traditional machine learning [7]. ResNet alleviates the vanishing gradient problem of deep networks through residual connections, further improving diagnostic accuracy. However, existing fault diagnosis methods based on CNN and ResNet cannot fully meet the multi-scale feature requirements of rolling bearing time-frequency images. Most CNN-based methods use a single-path serial network structure with fixed convolution kernel sizes, or a single-scale variation as layers deepen. This design cannot balance two aspects in time-frequency images: local fine textures that correspond to fault impacts, and global structural features that reflect the overall fault distribution. This leads to incomplete feature representation and difficulty in accurately capturing weak fault information [8]. Although ResNet solves the deep gradient vanishing problem, its convolution kernel design still follows a single-path serial structure, lacking the ability to capture information at different scales in parallel at the same level. Moreover, deep pooling operations tend to lose local fine fault features, resulting in insufficient diagnostic robustness under complex operating conditions such as variable speeds and strong noise [9]. The multi-scale convolutional neural network (MCNN) captures both local and global features at the same level by using parallel convolutional kernels of different sizes. It has demonstrated significant advantages in fields such as image classification and object detection. Lv et al. [10] proposed an intelligent fault diagnosis model based on MCNN and decision fusion, which can achieve a fault diagnosis accuracy rate of 99.8%. Li et al. [11] proposed an online transfer learning network based on MCNN for rolling bearing fault diagnosis, which can combine the online update strategy driven by prediction errors to achieve adaptive diagnosis of new data.

To further enhance the accuracy and generalization capability of fault diagnosis, classifiers such as SVM and ELM are widely used in the classification stage of bearing fault diagnosis, serving as an important bridge between deep learning feature extraction and fault classification. As an improved form of extreme learning machine (ELM), KELM introduces kernel function mapping to effectively solve the problems of poor classification

stability and insufficient generalization caused by random hidden layer parameters in ELM. It enables efficient classification of high-dimensional features without complex iterative training, offering advantages such as fast training speed and strong generalization [12]. Zhao et al. [13] addressed the problem of the remaining useful life prediction of rolling bearings by proposing a feature extraction method based on the maximum power spectral density curve derivative fitting method and constructed a prediction model combined with kernel extreme learning machine (KELM), verifying the superiority of the feature extraction and KELM fusion architecture in bearing condition assessment.

The classification performance of KELM heavily depends on the selection of its core hyperparameters, mainly the regularization coefficient and the kernel parameter. The regularization coefficient balances the model's fitting ability and generalization capability, while the kernel parameter determines the mapping effect of the feature space. The values of both directly affect the classification accuracy and stability of KELM [14]. In current research, various optimization algorithms have been introduced to determine the KELM parameters. For example, Wang et al. [15] systematically compared eight advanced metaheuristic optimization algorithms in displacement prediction tasks, confirming their effectiveness for tuning KELM parameters. The IGWO–KELM model proposed by Li et al. [16] adaptively optimized KELM by improving the grey wolf optimization algorithm, which not only improved classification accuracy but also eliminated redundant information through a feature selection strategy, significantly reducing computational cost. Zhong et al. [17] introduced a hierarchical RIME algorithm with multiple search preferences as an optimizer, providing reliable technical support for precise optimization of KELM hyperparameters and significantly improving the accuracy and robustness of bearing fault classification. In benchmark tests, RIME was systematically compared with 10 well-established algorithms and 10 of the latest improved algorithms. The results showed that RIME achieves better convergence accuracy and faster convergence speed, with an average accuracy improvement of about 28.5% over the 20 comparison algorithms.

In summary, this paper proposes a rolling bearing fault diagnosis method that integrates SWT, MCNN, and KELM optimized by RIME. First, SWT is used to transform vibration signals into time–frequency images with high energy concentration and clear ridges, effectively suppressing noise and highlighting fault impact features. Then, MCNN is employed to adaptively extract multi-scale features from the time–frequency images, compensating for the deficiency of traditional single-path serial networks in balancing local detail and global structural information. Finally, the RIME algorithm is introduced to adaptively optimize the regularization coefficient and kernel parameter of KELM, with KELM serving as the classifier. Benefiting from the flexible mapping capability of the RBF kernel, KELM can construct a more effective nonlinear decision boundary in the learned feature space, thereby achieving higher diagnostic accuracy than the traditional Softmax classifier.

The layout of this paper is as follows: Section 2 mainly introduces the detailed process of the bearing fault diagnosis method integrating the SWT and MCNN–RIME–KELM hybrid network model. Section 3 presents the theoretical foundations of the synchrosqueezed wavelet transform, multi-scale convolutional neural network, kernel extreme learning machine, and rime optimization algorithm. Section 4 presents experimental validations, where multiple comparisons are used to verify the effectiveness and superiority of the proposed method. Finally, some conclusions are drawn.

2. The Basic Process of Bearing Fault Diagnosis Methods

This paper proposes a rolling bearing fault diagnosis method integrating the SWT and MCNN–RIME–KELM hybrid network model. First, SWT is used to convert

one-dimensional non-stationary vibration signals into high-resolution time-frequency images, effectively suppressing energy dispersion and highlighting fault impact features, providing high-quality input for subsequent feature extraction. Second, the MCNN is constructed, which adaptively extracts multi-scale and multi-level deep fault features of local details and global structures from the time-frequency images through parallel branches with convolution kernels of different sizes. Finally, the RIME is introduced to adaptively optimize the regularization coefficient and kernel parameters of KELM, solving the problems of parameter sensitivity and low efficiency of manual parameter tuning, fully leveraging the classification advantages of KELM, and achieving high-precision fault classification. The detailed flow of the bearing fault diagnosis method is shown in Figure 1.

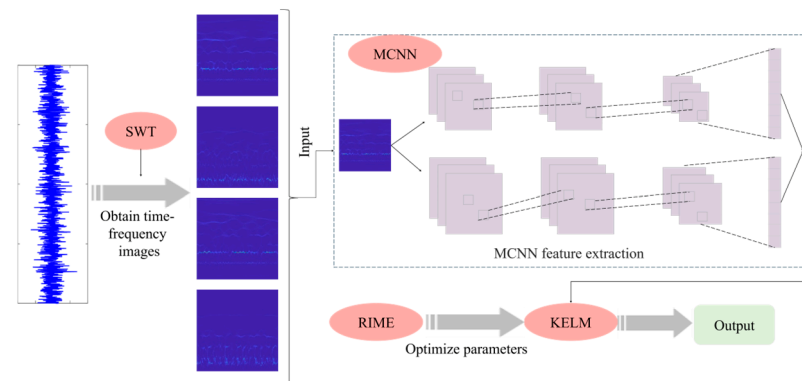


Figure 1. Flowchart of bearing fault diagnosis methods.

3. Methodology

3.1. SWT

The SWT, based on the continuous wavelet transform (CWT), reorders and compresses the wavelet coefficients along the frequency direction, significantly enhancing the time-frequency concentration. It can effectively highlight the impact components in the vibration signal and is suitable for extracting weak fault features. Its core idea is as follows: first, compute the CWT coefficients of the signal; then, estimate the instantaneous frequency (IF) at each time-frequency point; finally, compress and rearrange the wavelet coefficients along the frequency direction, transforming from the scale-time plane to the time-frequency plane, obtaining a highly energy-concentrated time-frequency representation, thereby achieving improved time-frequency resolution [18]. For a given signal $x(t)$, its CWT is defined as [3]

$$W_x(m, n) = \int_{-\infty}^{\infty} x(t) \frac{1}{\sqrt{m}} \psi^* \left(\frac{t-n}{m} \right) dt \quad (1)$$

where m is the scale parameter, n is the translation parameter, and $\psi(x)$ is the mother wavelet function (the complex Morlet wavelet is used in this paper), ψ^* denotes its complex conjugate. For any $W_x(m, n) \neq 0$, its instantaneous frequency can be calculated as

$$\omega_x(m, n) = -j \frac{1}{W_x(m, n)} \frac{\partial}{\partial n} W_x(m, n) \quad (2)$$

SWT transforms the CWT coefficients from the scale-time plane to the time-frequency plane and rearranges them along the frequency direction. This process can be expressed as

$$T_x(\omega_l, n) = \sum_{m_k: |\omega_x(m_k, n) - \omega_l| \leq \Delta\omega/2} W_x(m_k, n) m_k^{-3/2} \Delta m_k \quad (3)$$

SWT improves the energy concentration of the time–frequency representation, effectively suppresses energy dispersion, and makes the transient impact features caused by bearing faults more clearly identifiable in the time–frequency image.

3.2. MCNN

Traditional convolutional neural networks adopt a single–path serial structure, where the convolution kernel size is usually fixed or increases in a single scale as layers deepen. This design makes it difficult for the network to simultaneously capture local details and global structures in time–frequency images. Small convolution kernels can extract fine–grained, high–frequency texture information, but their limited receptive fields prevent them from perceiving the overall energy distribution. Large convolution kernels can cover larger receptive fields but tend to ignore local fine variations [19,20]. In contrast, MCNN constructs multiple branches using parallel convolution kernels of different sizes, extracting fine–grained details and macroscopic structural features from time–frequency images, and obtaining more comprehensive deep representations through feature fusion, thereby significantly enhancing feature extraction capability. Therefore, this paper adopts a dual–branch parallel multi–scale convolutional neural network to improve the accuracy of fault diagnosis [21]. The convolution kernel is the core part of convolution. By performing convolution operations with the convolution kernel and the input data matrix, it can effectively extract the local feature information of the data. The convolution operation can be expressed as

$$Y_i^l = f(W_i^l \otimes X^{l-1} + b_i^l) \quad (4)$$

where Y_i^l is the i –th feature in the l –th layer; W_i^l is the weight matrix of the i –th kernel in the l –th layer; $\otimes$ denotes the convolution operation; X^{l-1} is the output of the $l-1$ –th layer; and b_i^l is the bias term.

The proposed MCNN adopts a dual–branch parallel structure. Branch 1 uses larger convolution kernels to capture global contour features, while Branch 2 uses smaller kernels to extract local fine textures. The outputs of the two branches are fused by concatenation, followed by fully connected layers for classification, as shown in Table 1 below.

Table 1. The architecture of the dual–branch parallel MCNN.

Component	Layer Name	Kernel	Stride	Sizes
Input				(64, 64, 3)
Branch 1 (large kernel)	Conv1	7×7	1	(64, 64, 64)
Branch 1	MaxPool1	2×2	2	(32, 32, 64)
Branch 1	Conv2	5×5	1	(32, 32, 128)
Branch 1	MaxPool2	2×2	2	(16, 16, 128)
Branch 1	Conv3	3×3	1	(16, 16, 256)
Branch 1	GlobalAvgPool			(256)
Branch 2 (small kernel)	Conv1	3×3	1	(64, 64, 64)
Branch 2	MaxPool1	2×2	2	(32, 32, 64)
Branch 2	Conv2	3×3	1	(32, 32, 128)
Branch 2	MaxPool2	2×2	2	(16, 16, 128)
Branch 2	Conv3	3×3	1	(16, 16, 256)
Branch 2	GlobalAvgPool			(256)
Fusion	Concatenate			(512)
Fully Connected	FC1			(128)
Output	FC2			(num_class)

3.3. KELM

KELM introduces kernel function mapping based on the extreme learning machine (ELM), avoiding the instability caused by random hidden layer parameters and providing stronger generalization. Its performance is mainly determined by the regularization parameter C and the kernel parameter γ . ELM is a single-hidden-layer feedforward neural network whose input weights and hidden layer biases are randomly initialized, and the output weights are analytically computed via least squares, resulting in much faster training than traditional neural networks. However, the generalization performance of ELM is greatly affected by the number of hidden layer nodes [22]. KELM effectively overcomes this problem by replacing the random hidden layer mapping of ELM with a kernel function. The main idea of KELM is to replace the hidden layer output matrix $\mathbf{H}$ of ELM with the kernel matrix Ω_{ELM} according to Mercer's condition:

$$\Omega_{ELM} = \mathbf{H}\mathbf{H}^T, \Omega_{ELM,ij} = K(\mathbf{x}_i, \mathbf{x}_j) \quad (5)$$

where $K(\cdot, \cdot)$ is the kernel function. Kernel functions are stable and versatile. This paper adopts the Gaussian radial basis function (RBF) kernel:

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2\right), \gamma > 0 \quad (6)$$

where γ is the kernel parameter.

The output function of KELM is:

$$f(\mathbf{x}) = \begin{bmatrix} K(\mathbf{x}, \mathbf{x}_1) \\ \vdots \\ K(\mathbf{x}, \mathbf{x}_N) \end{bmatrix}^T \left(\frac{\mathbf{I}}{C} + \Omega_{ELM} \right)^{-1} \mathbf{T} \quad (7)$$

where C is the regularization coefficient, balancing model complexity and training error; $\mathbf{I}$ is the identity matrix; and $\mathbf{T}$ is the target label matrix of the training samples.

KELM inherits the advantages of fast training speed and good generalization from ELM while avoiding the problem of selecting the number of hidden layer nodes. The classification performance of KELM is highly dependent on the values of the regularization coefficient C and the kernel parameter γ ; improper parameter selection can lead to a significant decrease in classification accuracy.

3.4. RIME

Su et al. [23] proposed RIME, a novel optimization algorithm based on the physical mechanisms of rime ice formation and migration. It features strong global search capability, fast convergence speed, and a reduced tendency to fall into local optima. The algorithm consists of three core stages, which simulate the growth behavior of rime ice to achieve a dynamic balance between global exploration and local exploitation.

The first stage is the soft frost search stage, which simulates the random attachment of frost ice particles under light breeze conditions. In this stage, the position of each particle is updated as follows:

$$R_{ij}^{new} = R_{best,j} + r_1 \cdot \cos \theta \cdot \alpha \cdot (h \cdot (Ub_{ij} - Lb_{ij}) + Lb_{ij}), r_2 < E \quad (8)$$

where $R_{best,j}$ denotes the value of the current best individual in the j -th dimension; r_1 and r_2 are random numbers; θ is a factor controlling the granularity; α is a step function related to the ambient temperature; h is the adhesion parameter; Ub_{ij} and Lb_{ij} are the upper and

lower bounds of the search space in the j -th dimension, respectively; E is the adhesion coefficient, which can be expressed as

$$E = \sqrt{t/T} \quad (9)$$

where t and T denote the current iteration number and the maximum iteration number, respectively.

The second stage is the hard frost piercing stage, which simulates the intercrossing and stacking growth behavior among frost crystals. This mechanism performs dimension-wise crossover between ordinary agents and the optimal agent with a certain probability:

$$R_{ij}^{new} = R_{best,j} + r_2 \cdot \tan(\phi) \cdot (R_{best,j} - R_{ij}), r_3 < F_i \quad (10)$$

where r_3 is a random number, ϕ is the piercing angle, F_i is the normalized value of the current fitness. This mechanism facilitates information sharing within the population and effectively prevents the algorithm from getting trapped in local optima.

The third stage is forward greedy selection. It retains the current optimal solution and also introduces a suboptimal solution to guide population evolution in each iteration, thereby enhancing population diversity. This enhances population diversity and effectively prevents the algorithm from falling prematurely into locally suboptimal parameter combinations. As a result, RIME is able to search for the globally optimal hyperparameter configuration with a higher probability, thereby improving classification accuracy and robustness. In this paper, the population size of RIME is set to 10, and the maximum number of iterations is set to 20. The KELM parameters are optimized by RIME within the range [1, 100]. The specific optimization process is illustrated in Figure 2.

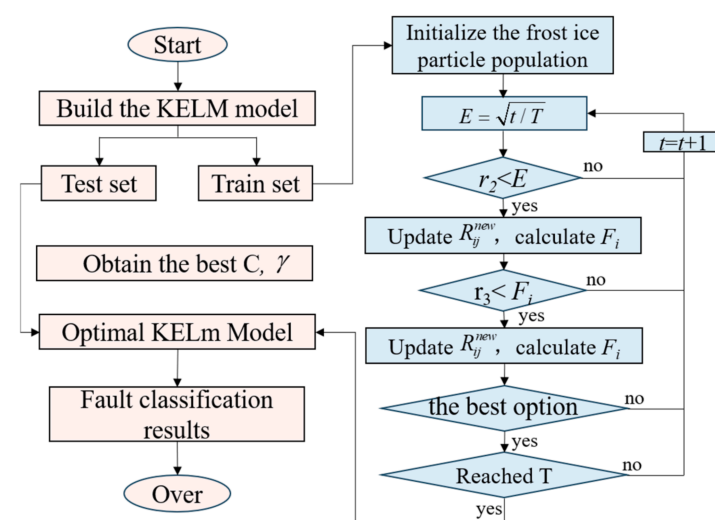


Figure 2. RIME optimizes the KELM process.

4. Experimental Validation

To validate the effectiveness of the proposed method, experiments are conducted on two public bearing datasets from Huazhong University of Science and Technology (HUST) and Case Western Reserve University (CWRU). The input features for all compared methods are uniformly represented by SWT time–frequency maps.

4.1. Huazhong University of Science and Technology Data Validation

The experiment uses the bearing fault dataset released by Huazhong University of Science and Technology [24]. This dataset was collected using a Spectra–Quest mechanical

fault simulator with ER–16K bearings. It covers nine condition categories, including normal state as well as different types of minor and severe faults. The signal acquisition system is equipped with a triaxial accelerometer and a tachometer. The experimental setup is shown in Figure 3.

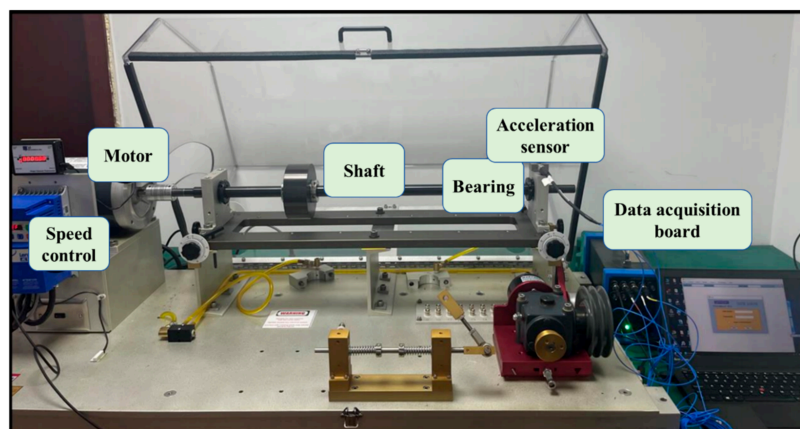


Figure 3. Bearing test rig at Huazhong University of Science and Technology.

The sampling frequency of the dataset is 25.6 kHz, and the data acquisition duration for each condition is 10.2 s. The collected data include three–directional acceleration signals as well as rotational speed information recorded by a tachometer. From these data, sample segments are extracted with a length of 4096 data points each, and an overlap of 0.1 s is set between consecutive segments. For each operating condition, 100 samples are collected from each of the three directional acceleration signals, resulting in 300 samples per condition. The dataset covers inner race faults, outer race faults, rolling element faults, and compound faults, each of which includes both minor and severe degrees, resulting in a total of 9 different operating conditions and 2700 samples. Among them, 70% are used as the training set and 30% as the test set, as detailed in Table 2.

Table 2. Dataset partition.

Fault Category	Label	Samples
Moderate Ball Fault	1	300
Moderate Compound Fault	2	300
Moderate Inner Race Fault	3	300
Moderate Outer Race Fault	4	300
Severe Ball Fault	5	300
Severe Compound Fault	6	300
Normal	7	300
Severe Inner Race Fault	8	300
Severe Outer Race Fault	9	300

Each segmented sample is converted into its corresponding SWT time–frequency representation, thereby highlighting fault characteristics in the time–frequency distribution. The time–domain signal of the bearing and its time–frequency map are shown in Figure 4.

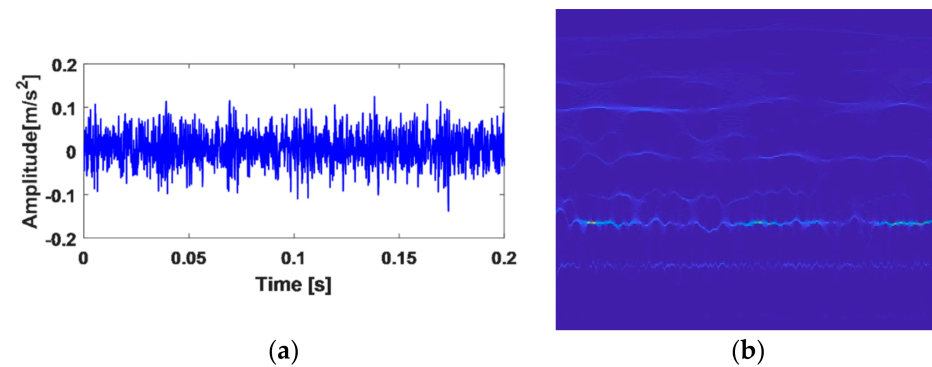


Figure 4. Bearing vibration signal: (a) time–domain signal; (b) SWT time–frequency representation.

The constructed time–frequency maps are fed into the established MCNN model, with the initial learning rate set to 0.001, the regularization coefficient set to 0.0001, and the Adam optimizer adopted. After feature extraction by the MCNN, the extracted features are imported into the KELM for classification, while the RIME algorithm is used to optimize its parameters. The final fault classification results are shown in Figure 5.

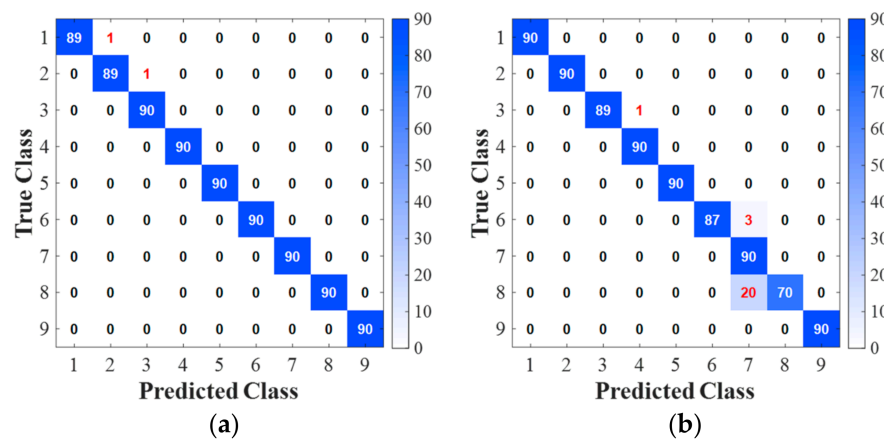


Figure 5. Confusion matrices of different methods (blue diagonal elements denote correct classifications, red off-diagonal elements denote misclassifications): (a) proposed method; (b) SWT + MCNN.

Figure 5 presents the confusion matrices for fault classification on the test set. Specifically, Figure 5a shows the classification results of the proposed method, achieving a high accuracy of 99.75%. In detail, only one fault of class 1 is misclassified as class 2, and one fault of class 2 is incorrectly assigned to class 3. In contrast, Figure 5b displays the results obtained by SWT + MCNN for classification, with an accuracy of only 97.04%. This method misclassifies a large number of class 8 faults into class 7, indicating that relying solely on the Softmax linear classifier yields unsatisfactory performance. The above comparison fully demonstrates that the proposed method has significant advantages in both feature extraction and classification decision –making, effectively enhancing the accuracy of bearing fault diagnosis.

Figure 6 visualizes the network classification results using t–SNE. Figure 6a shows the feature distribution at the initial input layer, and Figure 6b shows the feature distribution after MCNN–based feature extraction. It can be observed that, at the initial input stage, the features of different classes are heavily overlapped in the spatial distribution, and samples of the same class exhibit high dispersion, making effective discrimination directly from the raw input difficult. After feature extraction by the MCNN, however, samples of the same class become highly aggregated, while the boundaries between different classes are

clearly delineated. This indicates that the network has learned a highly discriminative deep feature representation, significantly improving class separability.

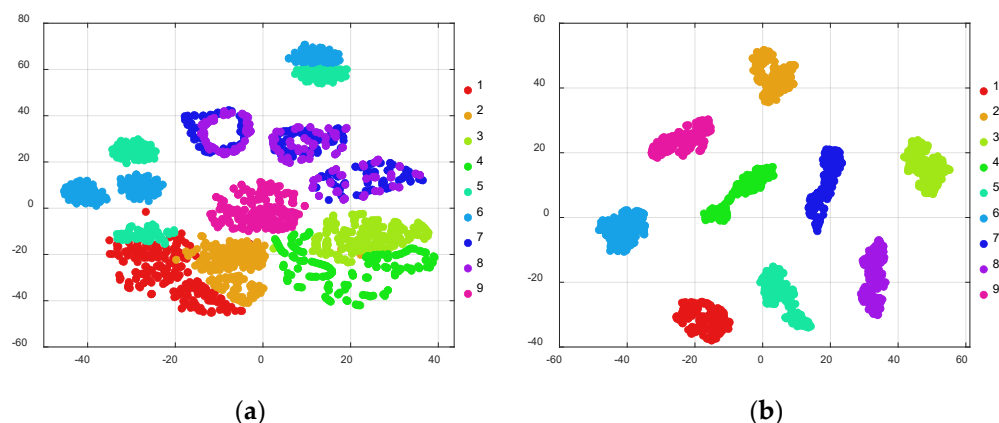


Figure 6. T–SNE visualization of the MCNN features: (a) initial input layer; (b) MCNN feature extraction.

Figure 7 presents the classification accuracies of five different methods under various SNR conditions. All results are reported as the mean of five repeated experiments, accompanied by error bars. When examining the classification performance of the five methods at an SNR of 10 dB, the proposed method achieves an average accuracy of 99.81% with a very small standard deviation, demonstrating that under high SNR conditions, it not only attains high classification accuracy but also exhibits good model stability. In contrast, the unoptimized MCNN + KELM yields a slightly lower accuracy, indicating that parameter tuning with RIME can improve classification performance. Due to the inherent limitation of the Softmax layer, MCNN alone performs worse than the proposed method. Furthermore, compared with two common networks, ResNet and AlexNet, the proposed method achieves superior accuracy in all cases.

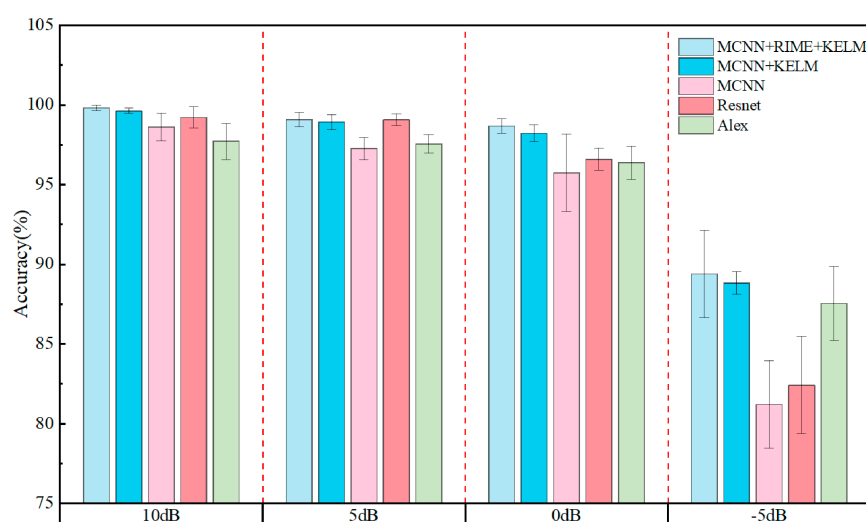


Figure 7. Method comparison under different SNRs.

Table 3 presents the specific classification accuracies of different methods under various SNR conditions. Comparing the results at 10 dB, 5 dB, 0 dB, and −5 dB reveals that as the SNR decreases, the accuracy of all methods declines to varying extents. The proposed method maintains high accuracies of 99.07% and 98.67% at 5 dB and 0 dB, respectively, with very small degradation, demonstrating strong anti-noise robustness. When the

SNR further drops to -5 dB, the performance of all methods deteriorates significantly; nevertheless, the proposed method still retains an accuracy of approximately 90%, while the other methods suffer a sharp drop, further validating its superior robustness under strong noise conditions. Moreover, judging from the trend of error bars, the proposed method maintains a consistently low standard deviation across all SNRs, indicating stable classification performance under different noise levels. By extracting fault features under different receptive fields via MCNN and then performing global optimization of KELM parameters using the RIME algorithm, the proposed method enables KELM to handle high-dimensional features with an optimal nonlinear classification boundary, thereby achieving high diagnostic accuracy under various SNR conditions.

Table 3. Accuracy of different methods under various signal–to–noise ratios.

	10 dB	5 dB	0 dB	−5 dB
MCNN + RIME + KELM	99.81 ± 0.17	99.07 ± 0.45	98.67 ± 0.45	89.38 ± 2.74
MCNN + KELM	99.63 ± 0.17	98.92 ± 0.48	98.21 ± 0.53	88.82 ± 0.72
MCNN	98.61 ± 0.87	97.26 ± 0.71	95.74 ± 2.43	81.20 ± 2.75
Resnet	99.21 ± 0.66	99.07 ± 0.35	96.59 ± 0.69	82.41 ± 3.04
Alex	95.96 ± 2.23	97.92 ± 0.59	96.37 ± 1.04	87.53 ± 2.31

4.2. Case Western Reserve University Data Validation

This experiment uses the publicly available bearing fault dataset from Case Western Reserve University [25]. The test platform mainly consists of a motor, a torque sensor, a dynamometer, and a control system. By using electrical discharge machining, single–point damages with diameters of 0.007 inch, 0.014 inch, and 0.021 inch are respectively introduced into the bearings, covering three types of faults: inner race, outer race, and rolling element. These faults result in nine different fault conditions, which together with the normal condition constitute a dataset of ten bearing operating states, covering different fault sizes of common bearing faults. Figure 8 illustrates the composition of the entire experimental setup.

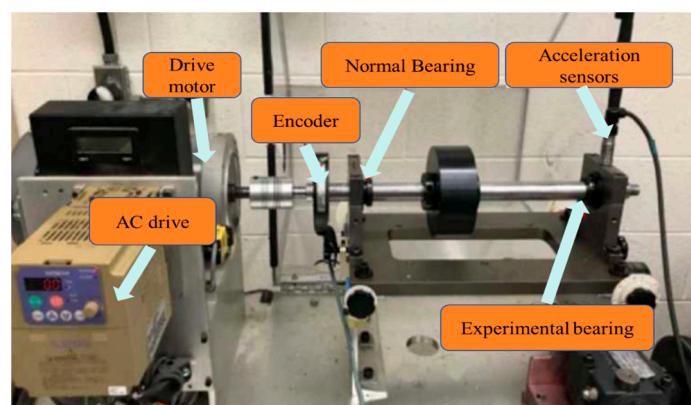


Figure 8. Case Western Reserve University Bearing Test Rig.

The bearing model used in the experiment is SKF6205, with a sampling frequency of 12,000 Hz. The duration of each data acquisition ranges from 10 to 20 s, from which 1024 data points are extracted to form one sample. For each fault category, 200 samples are generated, of which 140 are used for training and 60 for testing. The entire dataset consists of 2000 samples in total, with 1400 samples in the training set and 600 samples in the test set, as detailed in Table 4.

Table 4. Dataset partition.

Fault Category	Label	Samples
Normal	1	200
0.007 Inner Race Fault	2	200
0.007 Ball Fault	3	200
0.007 Outer Race Fault	4	200
0.014 Inner Race Fault	5	200
0.014 Ball fault	6	200
0.014 Outer race fault	7	200
0.021 Inner race fault	8	200
0.021 Ball fault	9	200
0.021 Outer race fault	10	200

Following the same processing procedure as in Section 3.1, the signal is first converted into an SWT time–frequency map, which is then fed into the MCNN network framework for feature extraction (with network parameters kept unchanged). Finally, classification is performed using the KELM with parameters optimized by RIME. The resulting classification outcomes are presented in Figure 9.

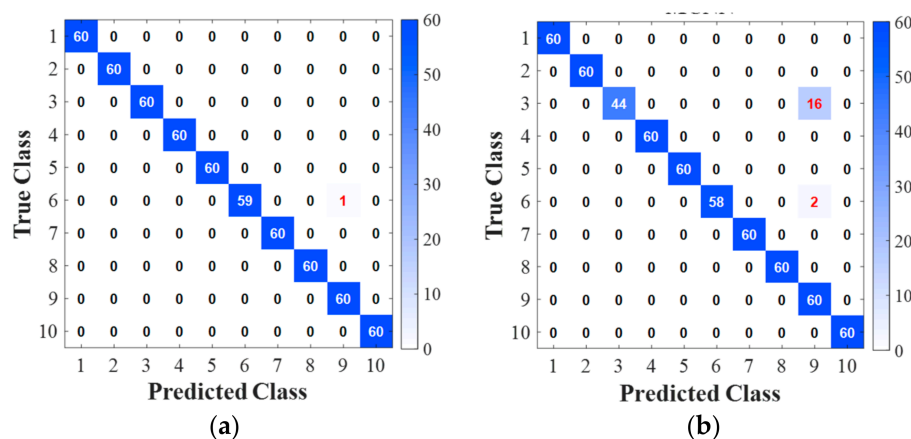


Figure 9. Confusion matrices of different methods (diagonal elements in blue denote correct classifications, off–diagonal elements in red denote misclassifications): (a) proposed method; (b) SWT + MCNN.

Figure 9 presents the confusion matrix for ten–class fault classification on this dataset. Specifically, Figure 9a shows the classification results of the proposed method, achieving an accuracy of 99.83%, with only one moderate rolling element fault being misclassified as a severe rolling element fault. Figure 9b displays the results obtained by SWT + MCNN, with an accuracy of only 97%, mainly characterized by the misclassification of many minor rolling element faults as severe ones. This indicates that when fault features are similar, the Softmax linear classifier performs significantly worse than the proposed method.

Figure 10 shows the t–SNE visualization of network classification on this dataset. Figure 10a presents the feature distribution at the initial input layer, and Figure 10b shows the feature distribution after MCNN–based feature extraction. It can be observed that the original features exhibit significant overlap and mixing in the spatial domain, with most categories intertwining in the central region, indicating a high degree of coupling among fault features in the original time–frequency representation. After deep mapping by the MCNN, all ten fault categories condense into compact and well–separated clusters, with samples within each cluster being highly concentrated. This fully demonstrates the

capability of the multi-scale convolutional structure to extract transient impulse features of faults, thereby providing highly discriminative features for the subsequent classifier.

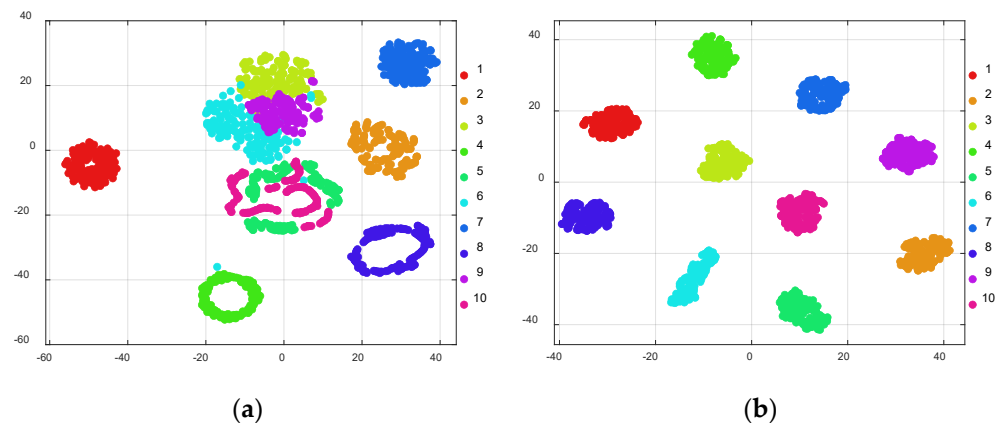


Figure 10. T-SNE visualization of the MCNN features: (a) initial input layer; (b) MCNN feature extraction.

Figure 11 presents a comparison of classification accuracies of five different methods under various signal-to-noise ratios. All results are the mean values of five repeated experiments, with error bars indicating the variation range. At an SNR of 10 dB, the proposed method achieves an average accuracy of 99.83% with a standard deviation of only 0.15%, demonstrating both high classification precision and good stability. The unoptimized MCNN + KELM attains an accuracy of 99.38%, slightly lower than that of the proposed method, indicating that RIME-based parameter optimization indeed improves classification performance. Constrained by the Softmax layer, the standalone MCNN yields an accuracy of 98.96%, which is inferior to the proposed method. Although the compared ResNet and AlexNet also achieve relatively high accuracies, both are lower than that of the proposed method and exhibit larger error bars.

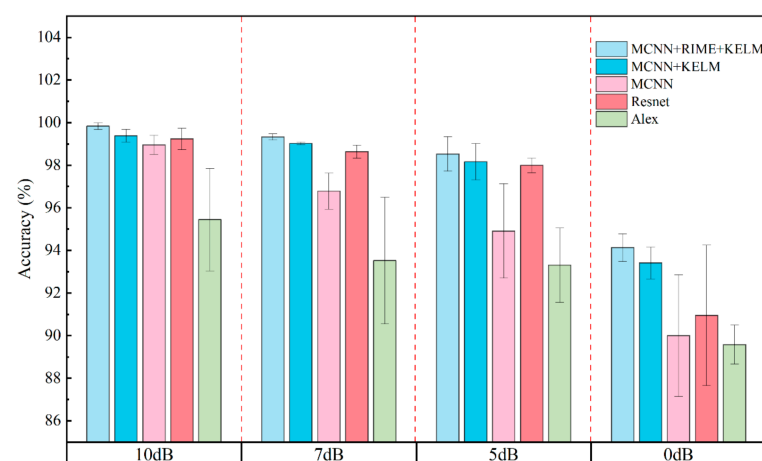


Figure 11. Classification results of different methods under different SNRs.

Table 5 lists the specific accuracies of different methods under various signal-to-noise ratios. At 10 dB and 5 dB, the proposed method still maintains an accuracy of approximately 99%, demonstrating good anti-noise capability. When the SNR drops to 0 dB, the proposed method achieves an accuracy of 94.14%, while MCNN declines to 90.00%, and ResNet and AlexNet reach only 90.95% and 89.58%, respectively. Moreover, the error bars of the latter three methods increase significantly, indicating severe fluctuations and insufficient reliability under strong noise conditions. Under all noise conditions, the proposed method

outperforms the compared methods in classification accuracy, and its error fluctuation range remains consistently low.

Table 5. Accuracy of different methods under various signal–to–noise ratios.

	10 dB	7 dB	5 dB	0 dB
MCNN + RIME + KELM	99.83 ± 0.15	99.33 ± 0.15	98.53 ± 0.81	94.14 ± 0.65
MCNN + KELM	99.38 ± 0.30	99.03 ± 0.07	98.17 ± 0.85	93.42 ± 0.75
MCNN	98.96 ± 0.44	96.79 ± 0.85	94.91 ± 2.21	90.00 ± 2.86
Resnet	99.24 ± 0.50	98.64 ± 0.31	98.00 ± 0.35	90.95 ± 3.30
Alex	95.45 ± 2.40	93.52 ± 2.97	93.31 ± 1.74	89.58 ± 0.91

5. Conclusions

To address the challenges of difficulty in fault feature extraction under strong background noise, limited classification capability of linear classifiers, and difficulty in adaptively optimizing classifier parameters for rolling bearings, this paper proposes a fault diagnosis method based on SWT combined with MCNN–RIME–KELM, achieving high–precision fault diagnosis of rolling bearings. The following conclusions are drawn:

(1) To overcome the limited classification capability of linear classifiers in convolutional networks, the KELM is introduced to replace the traditional Softmax linear classifier. Combined with the MCNN for extracting local texture and global contour features from time–frequency maps, and using the RIME to adaptively optimize the regularization coefficient and kernel parameters of KELM, the MCNN–RIME–KELM fault diagnosis model is constructed. This model significantly improves the accuracy of fault classification and nonlinear decision–making capability, effectively overcoming the fuzzy linear classification boundary limitation of Softmax.

(2) The proposed model is validated on two bearing datasets from Huazhong University of Science and Technology and Case Western Reserve University. The results show that the proposed method achieves average diagnostic accuracies of 99.75% and 99.83%, respectively, outperforming comparison models such as MCNN, MCNN + KELM, ResNet, and AlexNet. Furthermore, noise robustness experiments under different signal–to–noise ratios further demonstrate that the proposed method achieves the highest classification accuracy at all noise levels, with small error fluctuations in most cases, fully verifying the high precision and strong robustness of the algorithm.

Although the proposed method achieves favorable diagnostic performance, certain limitations still exist. Firstly, this study focuses primarily on bearing fault diagnosis under constant operating conditions; the adaptability of the model to more complex scenarios, such as variable loads and cross–device conditions, requires further validation. Secondly, the RIME optimization process involves iterative computations, which incur high time costs when handling large–scale industrial datasets. Therefore, future research will focus on improving the model’s generalization capability across different devices and operating conditions, lightweighting the model to enhance computational efficiency, and conducting detailed analyses of parameter sensitivity.

Author Contributions: L.W.: conceptualization, data curation, formal analysis, investigation, methodology, project administration, resources, supervision, validation, visualization, writing—original draft, writing—review and editing. X.L.: data curation, formal analysis, software, visualization, writing—review and editing. X.S.: data curation, investigation, resources, writing—review and editing. D.Z.: conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, project administration, resources, supervision, writing—original draft, writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Science and Technology Tackling Project of Henan Province, grant number 262102211047.

Data Availability Statement: No new data were created or analyzed in this study. The data used in this paper are available upon reasonable request.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

STFT	Short–Time Fourier Transform
CWT	Continuous Wavelet Transform
SWT	Synchrosqueezed Wavelet Transform
CNN	Convolutional Neural Networks
MCNN	Multi–scale Convolutional Neural Network
ELM	Extreme Learning Machine
KELM	Kernel Extreme Learning Machine

References

- Li, C.; Wang, Y.; Zhang, G.; Qin, Y.; Tang, B. An enhanced instantaneous angular speed estimation method by multi–harmonic time–frequency realignment for wind turbine gearbox fault diagnosis. *Meas. Sci. Technol.* **2023**, *34*, 085116. [\[CrossRef\]](#)
- Daubechies, I.; Lu, J.; Wu, H.-T. Synchrosqueezed Wavelet Transforms: An Empirical Mode Decomposition–like Tool. *Appl. Comput. Harmon. Anal.* **2011**, *30*, 243–261. [\[CrossRef\]](#)
- Wei, D.; Shen, J. Multi–Spectra Synchrosqueezing Transform. *Signal Process.* **2023**, *207*, 108940. [\[CrossRef\]](#)
- Luo, C.; Zong, Z. The Synchroextracting Algorithm Based on W Transform and Its Application in Channel Characterization. *IEEE Geosci. Remote Sens. Lett.* **2023**, *20*, 7502005. [\[CrossRef\]](#)
- Bao, W.; Tu, X.; Li, F.; Huang, Y. Generalized Synchrosqueezing Transform: Algorithm and Applications. *IEEE Trans. Instrum. Meas.* **2023**, *72*, 3503511. [\[CrossRef\]](#)
- Borghesani, P.; Ricci, R.; Chatterton, S.; Pennacchi, P. A New Procedure for Using Envelope Analysis for Rolling Element Bearing Diagnostics in Variable Operating Conditions. *Mech. Syst. Signal Process.* **2013**, *38*, 23–35. [\[CrossRef\]](#)
- Hou, J.; Jiao, H.; Zhong, Y.; Qi, J. An Intelligent Variable–Speed Bearing Fault Diagnosis Method Based on Order–Frequency Image Processing and Visual Transformer. *J. Mech. Sci. Technol.* **2026**, *40*, 2493–2502. [\[CrossRef\]](#)
- Deng, L.; Zhao, C.; Wang, X.; Wang, G.; Qiu, R. MRNet: Rolling Bearing Fault Diagnosis in Noisy Environment Based on Multi–Scale Residual Convolutional Network. *Meas. Sci. Technol.* **2024**, *35*, 126136. [\[CrossRef\]](#)
- Zhang, X.; Lv, J.; Wang, F.; He, F. A Dual–Channel Multi–Scale Residual Network with Dilated Convolution Optimization for Rolling Bearing Fault Diagnosis. *J. Vib. Control* **2025**, 10775463251403372. [\[CrossRef\]](#)
- Lv, D.; Wang, H.; Che, C. Multiscale Convolutional Neural Network and Decision Fusion for Rolling Bearing Fault Diagnosis. *Ind. Lubr. Tribol.* **2021**, *73*, 516–522. [\[CrossRef\]](#)
- Li, O.; Zhu, J.; Chen, M. Rolling Bearing Fault Diagnosis Based on Efficient Time Channel Attention Optimized Deep Multi–Scale Convolutional Neural Networks. *Meas. Sci. Technol.* **2024**, *35*, 126133. [\[CrossRef\]](#)
- Zhang, W.; Qi, R.; Ge, X.; Yang, G.; Bai, Y.; Yang, A. HOA–KELM: An Intelligent Diagnosis Method for Hot–Rolled Strip Manufacturing in the Industrial Internet of Things. *IEEE Internet Things J.* **2025**, *12*, 32344–32357. [\[CrossRef\]](#)
- Zhao, H.; Liu, H.; Jin, Y.; Dang, X.; Deng, W. Feature Extraction for Data–Driven Remaining Useful Life Prediction of Rolling Bearings. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 3511910. [\[CrossRef\]](#)
- Cao, L.; Yue, Y.; Zhang, Y. A Novel Fault Diagnosis Strategy for Heterogeneous Wireless Sensor Networks. *J. Sens.* **2021**, 2021, 6650256. [\[CrossRef\]](#)
- Wang, Y.; Sun, X.; Wen, T.; Wang, L. Step–like Displacement Prediction of Reservoir Landslides Based on a Metaheuristic–Optimized KELM: A Comparative Study. *Bull. Eng. Geol. Environ.* **2024**, *83*, 322. [\[CrossRef\]](#)
- Li, Q.; Chen, H.; Huang, H.; Zhao, X.; Cai, Z.; Tong, C.; Liu, W.; Tian, X. An Enhanced Grey Wolf Optimization Based Feature Selection Wrapped Kernel Extreme Learning Machine for Medical Diagnosis. *Comput. Math. Methods Med.* **2017**, 2017, 9512741. [\[CrossRef\]](#) [\[PubMed\]](#)
- Zhong, R.; Zhang, C.; Yu, J. Hierarchical RIME Algorithm with Multiple Search Preferences for Extreme Learning Machine Training. *Alex. Eng. J.* **2025**, *110*, 77–98. [\[CrossRef\]](#)

18. Ren, S.; Lou, X. Rolling Bearing Fault Diagnosis Method Based on SWT and Improved Vision Transformer. *Sensors* **2025**, *25*, 2090. [[CrossRef](#)] [[PubMed](#)]
19. Wu, Y.; Du, S.; Wu, G.; Guo, X.; Wu, J.; Zhao, R.; Ma, C. Minimum Maximum Regularized Multiscale Convolutional Neural Network and Its Application in Intelligent Fault Diagnosis of Rotary Machines. *ISA Trans.* **2025**, *159*, 1–21. [[CrossRef](#)] [[PubMed](#)]
20. Frannita, E.L.; Prananda, A.R. Mobile App–Based Leather Defects Identification with Fine–Tuned CNNs. *Measurement* **2026**, *259*, 119650. [[CrossRef](#)]
21. Yao, Z.; Jiang, Q.; Gu, X. A Fusion Model Fault Diagnosis Scheme Based on Multibranch Multiscale Residual Convolutional Network with Transformer. *Chem. Eng. Res. Des.* **2025**, *220*, 667–681. [[CrossRef](#)]
22. Huang, G.–B.; Zhou, H.; Ding, X.; Zhang, R. Extreme Learning Machine for Regression and Multiclass Classification. *IEEE Trans. Syst. Man. Cybern. Part B (Cybern.)* **2012**, *42*, 513–529. [[CrossRef](#)] [[PubMed](#)]
23. Su, H.; Zhao, D.; Heidari, A.A.; Liu, L.; Zhang, X.; Mafarja, M.; Chen, H. RIME: A Physics–Based Optimization. *Neurocomputing* **2023**, *532*, 183–214. [[CrossRef](#)]
24. Zhao, C.; Zio, E.; Shen, W. Domain Generalization for Cross–Domain Fault Diagnosis: An Application–Oriented Perspective and a Benchmark Study. *Reliab. Eng. Syst. Saf.* **2024**, *245*, 109964. [[CrossRef](#)]
25. Smith, W.A.; Randall, R.B. Rolling Element Bearing Diagnostics Using the Case Western Reserve University Data: A Benchmark Study. *Mech. Syst. Signal Process.* **2015**, *64–65*, 100–131. [[CrossRef](#)]

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.