Article citation info:
Li S. A rolling bearing fault diagnosis method based on BO-Caps net. Diagnostyka. 2026;27(2):2026211.
https://doi.org/10.29354/diag/224876.

DIAGNOSTYKA, 2026, Vol. 27, No. 2

e-ISSN 2449-5220
DOI: 10.29354/diag/224876

A ROLLING BEARING FAULT DIAGNOSIS METHOD BASED ON BO-CAPS NET

Shanshan LI *

School of Intelligent Manufacturing and Equipment, Lanzhou Vocational Technical College, Lanzhou, 730070, China
* Corresponding author, e-mail: 18809319966@163.com

Abstract

Addressing the common challenges of feature space relationship loss and reliance on tedious manual parameter tuning in deep learning for rolling bearing fault diagnosis, this paper proposes a fault diagnosis method grounded on Bayesian optimized capsule networks. First, the standard capsule network is optimized by designing multi-scale convolutional modules to capture fault information under different receptive fields. Then, a channel attention mechanism is introduced to dynamically weight the extracted multi-scale features. Furthermore, a Bayesian optimization framework is introduced to globally optimize the complex hyperparameter space of the improved model. Findings denote that the enhanced capsule network model achieves an accuracy of 99.8% under standard operating conditions, with its accuracy rapidly increasing to over 90% after 20 iterations. The proposed rolling bearing fault diagnosis method achieves an average F1 score of 99.7% in 10 repeated experiments, and only requires an average of 24 iterations to complete one hyperparameter optimization, verifying the comprehensive superiority of the method. The designed method, through the deep integration of innovative model structure and intelligent parameter optimization, effectively improves the accuracy and efficiency of diagnosis, providing a new paradigm for predictive maintenance of critical equipment under complex operating conditions.

Keywords: rolling bearing; fault diagnosis; capsule network; Bayesian optimization; multi-scale attention

1. BACKGROUND

In the "Industry 4.0" and intelligent manufacturing context, large rotating machinery is developing towards high speed, high precision, and heavy load, which poses a severe challenge to its operational reliability and safety [1-3]. As an essential component of such equipment, the condition of Rolling Bearings (RBs) directly affects the stable operation of the entire mechanical system [4-5]. According to statistics, among various failures of rotating machinery, downtime accidents caused by RB failure account for more than 40%. Especially in key fields such as wind power generation and high-speed railways, bearings operate under complex working conditions and strong noise for a long time. Once a sudden failure occurs, it may lead to catastrophic safety accidents [6]. Therefore, the urgent problem that the industry needs to solve is developing intelligent Bearing Fault Diagnosis (BFD) and realizing predictive maintenance.

In recent years, intelligent diagnostic methods for RB fault signals have become a research hotspot and core issue in the area of mechanical health monitoring. Zhang Q's team designed an intelligent Fault Diagnosis (FD) method grounded on Convolutional Neural Network (CNN) to address the issue that CNNs are not obvious in BFD. Outcomes denoted that the recognition accuracy of this method on the dataset was 99.96% [7]. Liu et al. developed an anti-noise BFD method grounded on improved recursive graph and CNN, to address the defect of recursive graph method being susceptible to noise interference in BFD. Outcomes denoted that the accuracy of this method could reach 90% under Gaussian white noise with a SNR greater than 6dB, verifying its effectiveness [8]. Zhao H's team designed a new parameter time redistribution multi-synchronous compression transform method to address the problem of non-stationary signal analysis generated during RB faults. The outcomes showed that this method had a high concentration of time-frequency energy under various noise levels and working conditions, and could accurately diagnose bearing faults [9]. To address the problem of difficulty in processing small sample data in rolling BFD, Zhao et al. designed an FD method grounded on continuous wavelet transform and connected neural network. Outcomes denoted that when the sample size was greater than 120, the

Received 2026-01-09; Accepted 2026-06-29; Available online 2026-06-29

accuracy could reach more than 99%, and it had good generalization performance under new categories and new working conditions [10].

The intelligent FD framework based on improved Capsule Networks (CapsNet) has been broadly utilized to enhance the fault identification performance under complex working conditions. Zhang et al. designed an improved CapsNet diagnostic model to deal with the issues of environmental noise interference and lack of massive labeled data in BFD. Outcomes denoted that the model improved the classification accuracy of BFD in noisy environments and the training process was fast and efficient [11]. Shi X team proposed a diagnostic method combining domain knowledge and CapsNet to address the problem of different dissolved gas concentrations in oil under different operating conditions in power transformer FD. Outcomes denoted that the diagnostic accuracy of this method on real databases was as high as 94.02% [12]. Wang et al. designed a multi-scale spatiotemporal CapsNet model to tackle the issue of effectively capturing multi-scale and time series features in BFD. The model achieved an average accuracy of over 97% in noise-free environments and over 80% in noisy environments [13]. Chai et al. proposed an improved CapsNet for small sample BFD to address the problems of insufficient fault signals and difficulty in extracting sensitive fault features in intelligent FD. The results showed that the method could still accurately diagnose faults in small sample and variable noise tests, which is better than some advanced deep learning methods [14].

In summary, while progress has been made in the research of methods for rolling BFD, two key shortcomings remain: Firstly, the ability of models to extract features is limited when they are faced with non-stationary fault signals under complex operating conditions, which makes it difficult to capture weak features effectively. Secondly, the performance of deep models depends heavily on manually configuring hyperparameters, which makes the optimization process cumbersome. Therefore, to solve these shortcomings, this study designs an FD method based on Bayesian Optimization-based Capsule Network (BO-CapsNet), aiming to construct a new paradigm for intelligent diagnosis of RBs that combines high accuracy and high automation, providing reliable technical support for ensuring the safe and stable operation of critical mechanical equipment. The innovation of this research lies in building upon CapsNet by integrating multi-scale convolution and Channel Attention Mechanisms (CAMs) to construct an improved model with stronger feature representation capabilities. Simultaneously, it introduces the Bayesian Optimization (BO) framework, achieving deep coupling between model structure optimization and intelligent parameter configuration. This study proposes a novel paradigm that combines high precision and efficiency for addressing intelligent diagnosis of rolling bearings under complex operating conditions. It particularly leverages the explicit modeling capability of capsule networks to handle the prevalent non-stationarity, multi-scale characteristics, and localized impact features in bearing vibration signals, thereby providing a theoretical foundation for the continued application of capsule networks in industrial fault diagnosis amid rapid advancements in sequence modeling architectures.

2. METHODS

To deal with the issues of insufficient feature extraction and difficulties in model parameter optimization in rolling BFD, this study develops an FD method based on BO-CapsNet. First, an improved CapsNet model integrating multi-scale convolution and CAMs is constructed to enhance its feature representation capabilities. Then, the BO framework is introduced to automatically optimize the model's hyperparameters, ultimately forming a complete diagnostic process from model structure optimization to intelligent parameter configuration.

2.1 Construction of an improved capsnet model incorporating multi-scale attention mechanisms

The problem of weak fault characteristics and low signal-to-noise ratio in RBs under complex working conditions, which makes diagnosis difficult [15-16]. It is greatly significant to study intelligent diagnostic methods that can deeply mine the intrinsic structural information of signals. Presently, diagnostic methodologies founded upon conventional signal processing and machine learning have been extensively utilized [17]. However, they extract mostly shallow features, which are difficult to fully characterize the nonlinearity of faults, resulting in limited diagnostic accuracy and generalization ability under complex working conditions. Therefore, this study introduces CapsNet, which performs well in characterizing the hierarchical relationship of features, as the basic model. CapsNet uses vector-form "capsule" units to encapsulate multiple attributes of features and preserves the spatial hierarchical relationship between features through a dynamic routing mechanism. The specific structure is denoted in Fig. 1.

In Fig. 1, the CapsNet model first receives the preprocessed vibration signal through an input layer. The signal then flows through convolutional layers to capture its long-range dependency features. Subsequently, two small convolutional kernel layers perform deeper local feature extraction. The extracted feature maps are fed into the main Capsule Layer (CL), where they are reshaped into primary capsule vectors. Finally, these primary capsules converge to the digital CL using a dynamic routing algorithm, and the output layer outputs the probability of each fault category. The primary capsules adaptively route information to the digital CL through CapsNet's core dynamic routing algorithm, achieving advanced feature aggregation and classification. Its operating principle is shown in Fig. 2.

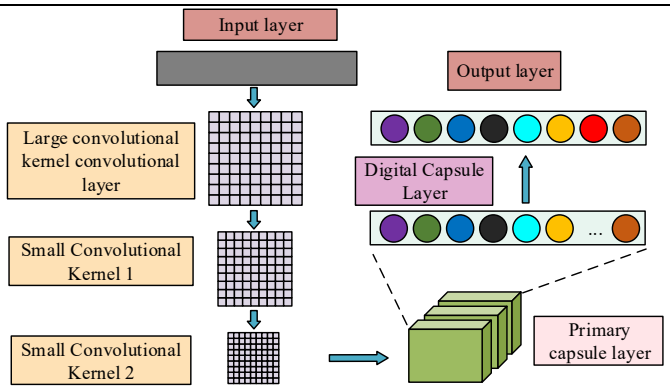


Fig. 1. CapsNet model architecture diagram

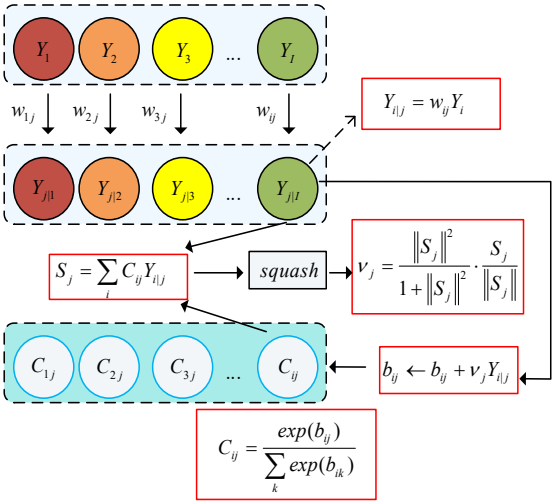


Fig. 2. Schematic diagram of dynamic routing algorithm operation

In Figure 2, the output of the main CL is passed to the digital CL through a dynamic routing algorithm unique to capsule networks. The algorithm dynamically determines the connection strength between lower-level capsules and higher-level capsules through an iterative "voting" mechanism. This process ensures that the spatial hierarchy of features can be maintained and passed on in the network. The algorithm first uses a transformation matrix w_{ij} to convert each lower-level capsule into a prediction vector $Y_{i|j}$. The expression of which is shown in Equation (1) [18].

$$Y_{i|j} = w_{ij} Y_i \tag{1}$$

In Equation (1), Y_i means the output vector of the i-th capsule in the main CL i. The total input vector S_j of the digital capsule j is obtained by weighted summation of $Y_{i|j}$, and its expression is denoted in Equation (2).

$$S_j = \sum_i C_{ij} Y_{i|j} \tag{2}$$

In Equation (2), C_{ij} means the coupling coefficient determined by the dynamic routing process. The specific expression for C_{ij} is shown in Equation (3).

$$C_{ij} = \frac{exp(b_{ij})}{\sum_k exp(b_{ik})} \tag{3}$$

In Equation (3), k represents the capsule index, and b_{ij} represents the number of routing pairs connecting the main capsule $i$and the digital capsule$j$. The specific expression for b_{ij} is shown in Equation (4).

$$b_{ij} \leftarrow b_{ij} + v_j Y_{i|j} \tag{4}$$

In Equation (4), $\leftarrow$ represents the assignment operation, and v_j represents the final output vector of the digital capsule j after activation, as shown in Equation (5).

$$v_j = \frac{\|s_j\|^2}{1+\|s_j\|^2} \cdot \frac{s_j}{\|s_j\|} \tag{5}$$

The direction of the capsule output vector is preserved through Equation (5), while ensuring that its magnitude is between (0,1). Capsule networks, through their unique dynamic routing mechanism, demonstrate advantages in preserving and utilizing the spatial hierarchy of features. However, the quality of the primary features input to the main CL has a significant impact on their overall performance. Standard CapsNet models typically employ only single-scale convolution operations at the front end, which limits the breadth of their receptive field and their ability to capture instantaneous impact details and long-range periodic patterns in RB vibration signals. Therefore, this study optimizes the feature extraction layer of standard CapsNet by introducing a multi-scale convolution module to compensate for this deficiency. This module uses parallel multi-size convolution kernels to extract and fuse multi-scale features, providing a comprehensive representation of the CL. A specific schematic diagram is shown in Figure 3.

In Figure 4, the input feature map information flow is split into two paths. One path is compressed using global average pooling to generate descriptors that represent global information for each channel. These descriptors are then fed into an activation module, which is a two-layer fully connected neural network. The neural network learns the correlations between channels and generates dynamic weights for each channel. Finally, the weight vector is multiplied by the original input feature map, channel by channel, to obtain the final output. In summary, this study first employs multi-scale convolution in parallel to extract fault features under different receptive fields, and secondly utilizes a CAM to dynamically weight the extracted multi-scale features. These two

improvements work synergistically to ultimately construct the proposed improved CapsNet model.

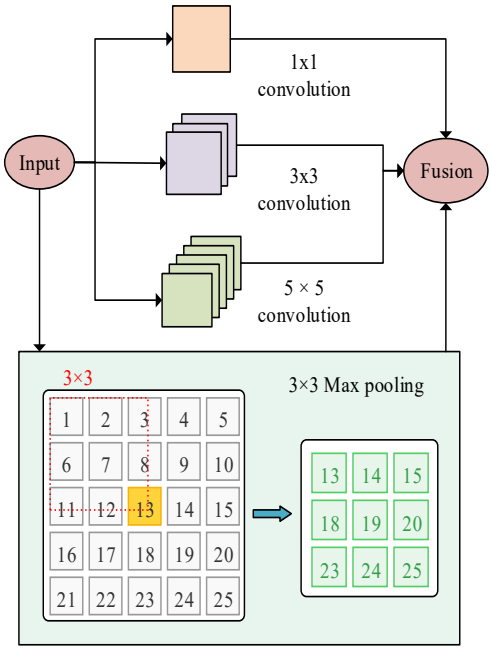


Fig. 3. Schematic diagram of multi-scale convolution module

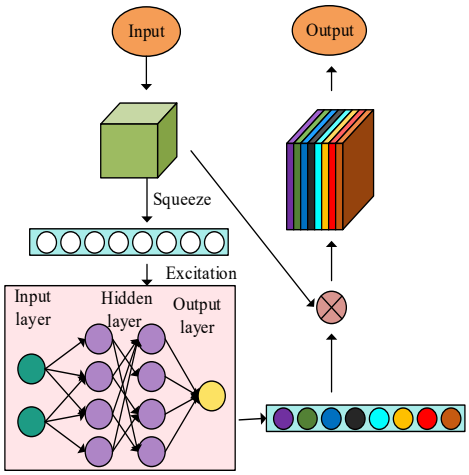


Fig. 4. Schematic diagram of the CAM

2.2 Design of a model hyperparameter optimization method based on BO

The enhanced CapsNet model effectively enhances the model's ability to extract relevant fault features from the original vibration signal by incorporating multi-scale convolution and attention mechanisms. However, in practical engineering applications, the improved model's high performance depends heavily on an optimal set of hyperparameter configurations. These parameters together constitute a complex high-dimensional space, and their selection directly affects the model's convergence speed and generalization ability [19]. If the configuration is not appropriate, even with an advanced model structure, it is difficult to achieve the ideal diagnostic effect. Therefore, this study introduces the BO framework to optimize the model's hyperparameters. This algorithm can find a set of parameter combinations that make the model's performance optimal within a limited number of evaluations. Its optimization objective is to find the optimal parameter combination $\widetilde{\theta}$ that maximizes the objective function $f(\theta)$. The specific expression is denoted in Equation (7).

$$\widetilde{\theta} = argmax_{\theta\in\Theta} f(\theta) \quad (7)$$

In Equation (7), θ represents any combination of hyperparameters, and Θ represents the search space consisting of all possible combinations of hyperparameters. To model $f(\theta)$, the study uses the Gaussian process $\mathcal{GP}$ in the BO framework as a surrogate model. The specific expression is shown in Equation (8) [20].

$$f(\theta) \sim \mathcal{GP}(m(\theta), \beta(\theta, \theta')) \quad (8)$$

In Equation (8), $m(\theta)$ represents the mean function, θ' represents another set of hyperparameter combinations, and $\beta(\theta, \theta')$ represents the kernel function. The expression for $\beta(\theta, \theta')$ is shown in Equation (9).

$$\beta(\theta_g, \theta_h) = \sigma_f^2 \left(1 + \sqrt{5}r + \frac{5}{3}r^2\right) exp(-\sqrt{5}r) \quad (9)$$

In Equation (9), σ_f^2 represents the signal variance, and r means the weighted Euclidean distance between the hyperparameter combinations θ_g and θ_h. The choice of $\beta(\theta, \theta')$ provides the prior structure for the Gaussian process. During the iteration process, each time the model is trained with θ_g, an observation data pair is obtained. After collecting t data pairs, the Gaussian process uses this information to update its understanding of $f(\theta)$. At the same time, it gives the posterior prediction distribution $\mu_t(\theta_*)$ for any new point θ_*, and its expression is shown in Equation (10).

$$\mu_t(\theta_*) = \beta_*^T (B + \sigma_n^2 I)^{-1} d \quad (10)$$

In Equation (10), β_* represents the kernel function vector between the new point θ_* and all observed points, B represents the covariance matrix between the observed points, σ_n^2 represents the noise variance, I means the identity matrix, and d means the function value vector of the observed points. The Gaussian process can also give the uncertainty of the prediction, i.e., the posterior variance $\sigma_t^2(\theta_*)$, which is expressed as shown in Equation (11) [21-22].

$$\sigma_t^2(\theta_*) = \beta(\theta_*, \theta_*) - \beta_*^T (B + \sigma_n^2 i)^{-1} \beta_* \quad (11)$$

In Equation (11), $\sigma_t^2(\theta_*)$ quantifies the degree of uncertainty in the prediction at θ_*. The surrogate model provides the prediction mean and variance, and the next step requires a collection function to guide the search direction. This study adopts expectation lifting as the collection function. First, a standardized variable Q is defined, and its expression is denoted in Equation (12).

$$Q = \frac{\mu_t(\theta) - d^+ - \xi}{\sigma_t(\theta)} \quad (12)$$

In Equation (12), d^+ represents the maximum value among the current observations, and ξ represents the tradeoff parameter. Based on this standardized variable, the complete expression for the expected improvement $EI(\theta)$ is shown in Equation (13).

$$EI(\theta) = (\mu_t(\theta) - d^+ - \xi)\Phi(Q) + \sigma_t(\theta)\phi(Q) \quad (13)$$

In Equation (13), Φ and ϕ represent the cumulative distribution function and the probability density function, respectively. In each iteration, the BO framework needs to find the next acquisition point that maximizes the acquisition function. Its expression is shown in Equation (14).

$$\theta_{t+1} = argmax\, EI(\theta) \tag{14}$$

In Equation (14), θ_{t+1} represents the combination of hyperparameters to be evaluated in the next iteration. The study applies the BO framework to improve the CapsNet model, forming an automated hyperparameter optimization framework. The specific details are shown in Figure 5.

In Figure 5, the process first randomly samples several sets of initial hyperparameters and evaluates their initial performance by training and validating the model, thereby constructing an initial observation dataset D. Then, the posterior distribution of the Gaussian process is updated using the current dataset D_t, and θ_{t+1} is determined using the $EI(\theta)$ function. The improved CapsNet model is validated based on θ_{t+1} to obtain the true performance value. New observation data pairs are added back to the dataset for the next round of Gaussian process updates. When the termination condition is met, the algorithm returns θ' until the end. The specific expression for D_t is shown in Equation (15).

$$D_t = D_{t-1} \cup \{(\theta_t, f(\theta_t))\} \tag{15}$$

In Equation (15), D_{t-1} means the dataset from the previous round, and $(\theta_t, f(\theta_t))$ represents the newly evaluated hyperparameters and their corresponding performance values. An automated optimization framework can be used to find the optimal hyperparameters for improving the CapsNet model. This study proposes a BO-CapsNet method for rolling BFD, and its complete implementation process is denoted in Figure 6.

In Figure 6, the proposed BO-CapsNet FD method begins with data preparation. Subsequently, the optimal hyperparameters are iteratively searched using the BO framework. The algorithm continuously updates the surrogate model and optimizes the acquisition function during iterations. Simultaneously, the performance of the improved CapsNet model is tested using the selected hyperparameters until the termination condition is met. After optimization, the final stage begins. First, the obtained optimal hyperparameters are used for training to construct the final diagnostic model. In the application stage, this method calls this model to process new test data and outputs the final FD result. In summary, this research constructs an improved CapsNet integrating multi-scale convolution and attention mechanisms and introduces the BO framework for automated hyperparameter optimization of this improved model. Finally, it deeply integrates the two to form the BO-CapsNet FD method, realizing a complete design process from model structure optimization to intelligent parameter configuration.

3. RESULTS AND ANALYSIS

The study designed two progressively comprehensive experiments. First, to assess the structural advantages of the improved CapsNet model, it was compared with other models. For evaluation metrics, accuracy, F1 score, and recall were used, and the dataset was divided. Next, BO-CapsNet was validated using metrics such as loss value and number of iterations.

3.1 Performance validation of improved CapsNet model

To test the performance advantages of the improved CapsNet model, comparative experiments were performed. The experimental data came from the internationally recognized Case Western Reserve University bearing dataset. Data covering normal operation and various typical failure modes were selected, and a total of 2500 sample points were constructed using sliding window sampling. All samples were broken into 1750 training samples and 750 test samples in a 7:3 ratio, and normalized before being input into the model. The experimental environment and key parameter settings are represented in Table 1.

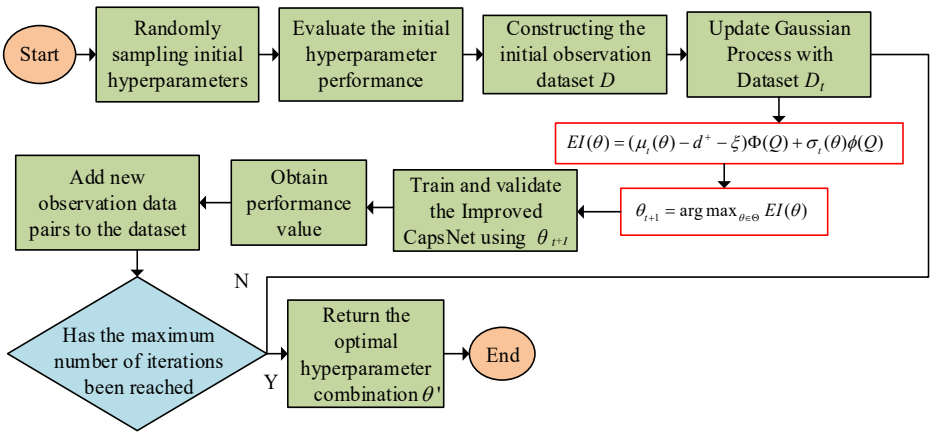


Fig. 5. BO framework optimizes Caps Net Flowchart

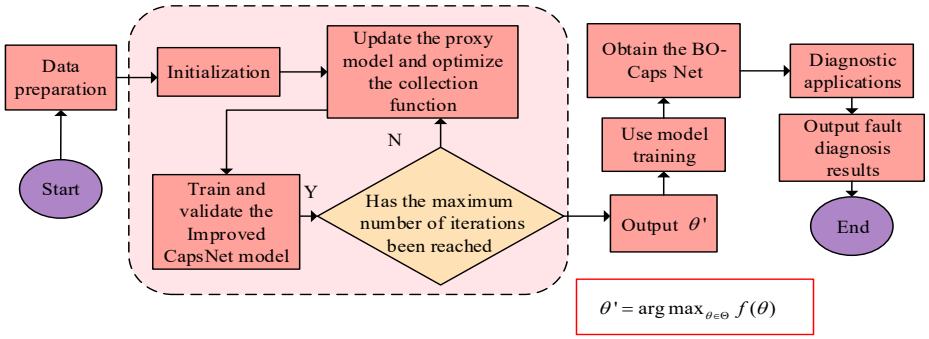


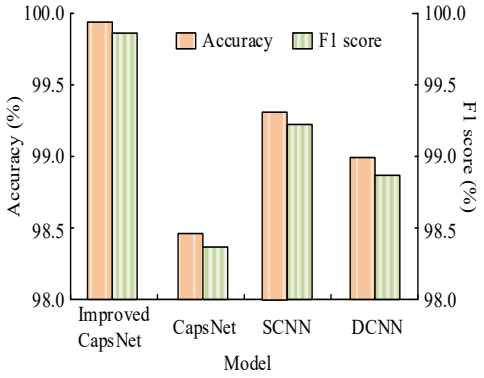
Fig. 6. Rolling BFD method based on BO-CapsNet

Table 1. Experimental environment and parameter settings

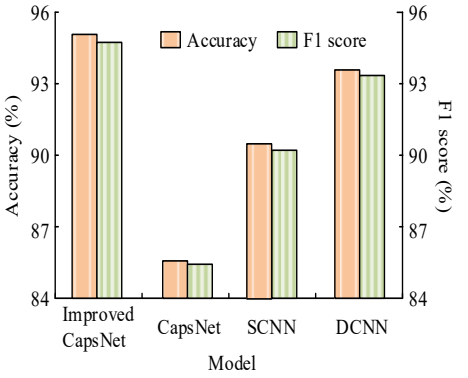
Experimental environment	Configuration information
Operating system	Windows 11
Memory	32 GB
CPU	Intel Core i7-12700H
GPU	NVIDIA GeForce RTX 3060 (8 GB)
Programming language	Python 3.9
Deep learning framework	PyTorch 1.12
Initial learning rate	0.001
Input data dimension	640×640
Number of training rounds	100
Batch size	64
Maximum iterations	100

This study compared the improved CapsNet model with other CapsNet, Spatial CNN (SCNN), and Deep CNN (DCNN) models under both standard and complex operating conditions. The accuracy and F1 score of each model under various conditions are denoted in Figure 7. In Figure 7(a), under the standard operating condition, the improved CapsNet model achieved an accuracy of 99.8% and an F1 score of 99.7%. The CapsNet model achieved an accuracy of 98.4% and an F1 score of 98.3%, respectively. In Figure 7(b), under the complex operating condition, the improved CapsNet model maintained a relatively high accuracy and F1 score of approximately 94.8% and 94.5%, respectively. In contrast, CapsNet experienced the most significant performance degradation, with its accuracy and F1 score dropping to approximately 85.8% and 85.6%, respectively. This demonstrates that the improved model exhibits stronger stability and generalization ability, validating the effectiveness of the improvement strategy.

The study compared the accuracy and loss values of various models, and the findings are denoted in Figure 8. Figure 8(a) showcases that the accuracy of the improved CapsNet model rapidly increased to over 90% after about 20 iterations, eventually stabilizing at approximately 94%. The accuracy of the CapsNet model eventually stabilized at approximately 90%. Figure 8(b) shows that the loss value of the improved CapsNet model decreased the fastest, dropping below 0.1 after 40 iterations,



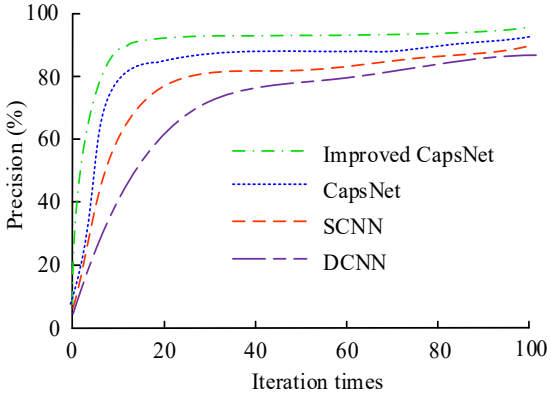
(a) Performance comparison of various models under standard operating conditions



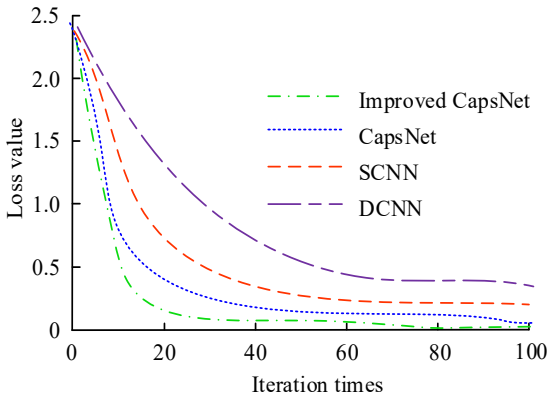
(b) Performance comparison of various models under complex operating conditions

Fig. 7. Comparison of model performance under different working conditions

and eventually converging to a minimum level close to 0.01. The loss values of the CapsNet and SCNN models also decreased effectively, eventually stabilizing at approximately 0.12 and 0.2, respectively. This reveals that the improved CapsNet model not only achieves optimal final performance but also has a significant advantage in training efficiency.



(a) Comparison of accuracy of various models

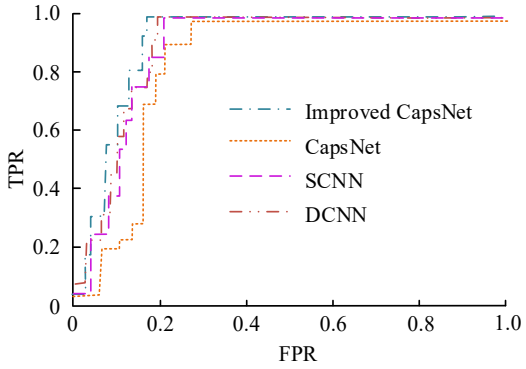


(b) Comparison of loss values for different models

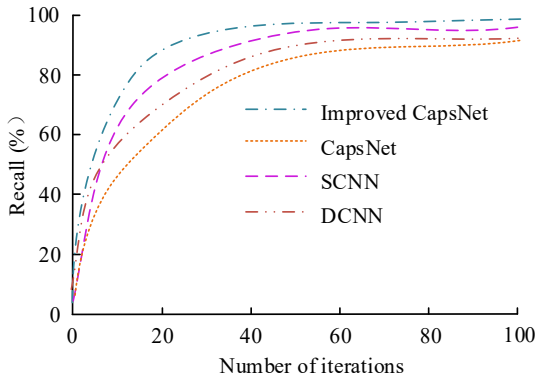
Fig. 8. Comparison of accuracy and loss values of different models during training

The study further validated the Receiver Operating Characteristic Curve (ROC) and recall of each model, as denoted in Figure 9. Figure 9(a) showcases that the improved CapsNet model achieved a True Positive Rate (TPR) of approximately 0.97 with a False Positive Rate (FPR) of 0.15. At the same level, the TPRs of SCNN, DCNN, and CapsNet were approximately 0.85, 0.80, and 0.70, respectively. Figure 9(b) shows that the improved CapsNet model exhibited the fastest recall growth, reaching nearly 95% after approximately 30 iterations and eventually stabilizing at a peak of over 98%. The SCNN model showed the second fastest convergence speed and final performance, eventually stabilizing at around 95%. The DCNN and CapsNet models ultimately stabilized at approximately 92% and 89% recall, respectively.

In recent years, fault diagnosis methods based on sequence modeling architectures such as long short-term memory networks, sequential convolutional networks, and Transformers have garnered



(a) Comparison of ROC curves of various models



(b) Comparison of recall rates for different models

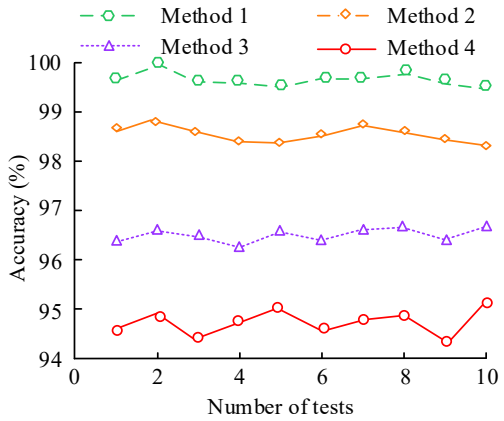
Fig. 9. Comparison of ROC curves and recall convergence processes across different models

significant attention. These approaches excel at capturing long-term temporal dependencies and demonstrate strong modeling capabilities for continuous stationary signals. However, rolling bearing vibration signals typically exhibit characteristics like transient impulses, local abrupt changes, and multi-scale frequency coupling, with critical fault information primarily emerging from the interplay between local structural patterns and features across different scales. In contrast, CapsNet represents feature attributes using vector capsules and employs a dynamic routing mechanism to explicitly model hierarchical relationships between features, thereby preserving more structural information during feature propagation while avoiding spatial loss associated with traditional convolutional pooling. Additionally, the introduced multi-scale convolutional modules enhance fault feature extraction across diverse sensing fields, while the channel attention mechanism enables adaptive reinforcement of key fault information. Under identical experimental conditions, the improved CapsNet not only maintains rapid convergence but also achieves diagnostic accuracy of 94.8% and recall rates exceeding 98% even under complex operating conditions, demonstrating superior feature representation performance for non-stationary rolling bearing fault signals. From an engineering application perspective, this model maintains high diagnostic precision while offering relatively manageable parameter complexity and lower requirements for training data volume and

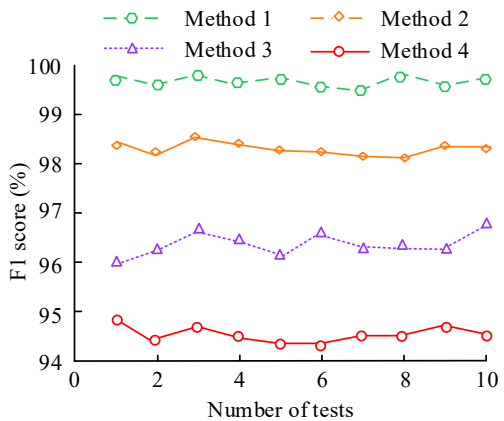
computational resources compared to large-scale sequence modeling architectures, making it particularly suitable for rapid deployment and online diagnostic tasks in industrial settings.

3.2. Comprehensive evaluation of BO-CapsNet FD method

This study compared four FD methods: BO-CapsNet (Method 1), Bayesian Optimization Deep CNN (BO-DCNN) (Method 2), Genetic Algorithm Capsule Network (GA-CapsNet) (Method 3), and Particle Swarm Optimization CNN (PSO-CNN) (Method 4). The accuracy and F1 score of each method in 10 tests are shown in Figure 10. Figure 10(a) shows that Method 1 had an average accuracy of 99.7%, while Method 2 had an average accuracy of 98.6%. Figure 10(b) shows that Method 1 had an average F1 score of 99.7%, while Method 4 had the lowest average F1 score of 94.6%. Overall, Method 1 demonstrates extremely high stability in repeated experiments.



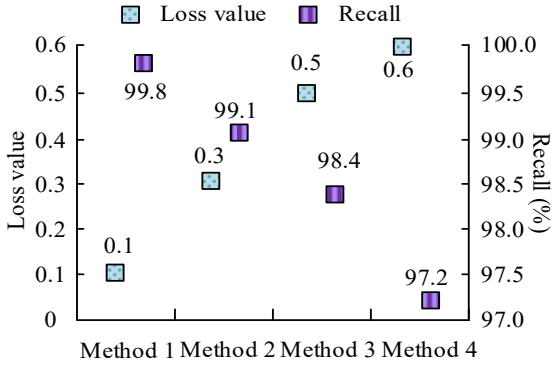
(a) Comparison of accuracy of different methods in 10 experiments



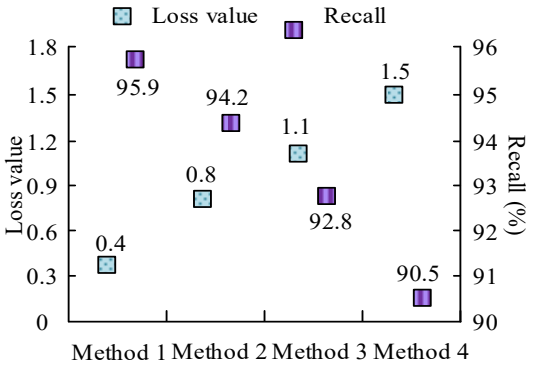
(b) Comparison of F1 scores of different methods in 10 experiments

Fig. 10. Comparison of performance stability of different FD methods

The study compared the loss values and recall rates of various methods under standard and complex operating conditions, and the findings are denoted in Figure 11. Figure 11(a) shows that under standard operating conditions, Method 1 had a loss value of only 0.1 and a recall rate of 99.8%. Method 2 had a loss value of 0.3 and a recall rate of 99.1%. Method 3 had a loss value of 0.5 and a recall rate of 98.4%. Method 4 had the weakest performance, with a loss value as high as 0.6 and a recall rate of 97.2%. Figure 11(b) shows that although the loss value of Method 1 increased to 0.4, it was still the lowest among the four methods, and its recall rate remained at a relatively high level of 95.9%. The loss values of Methods 2, 3, and 4 increased dramatically to 0.8, 1.1, and 1.5, respectively, while the recall rates decreased to 94.2%, 92.8%, and 90.5%, respectively. Therefore, the proposed method achieves the best fitting effect and the highest fault identification rate.



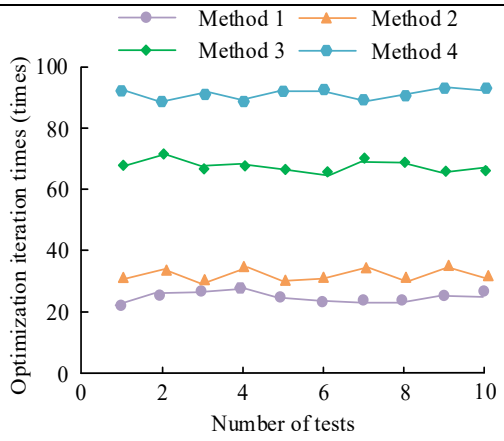
(a) Performance comparison of different methods under standard operating conditions



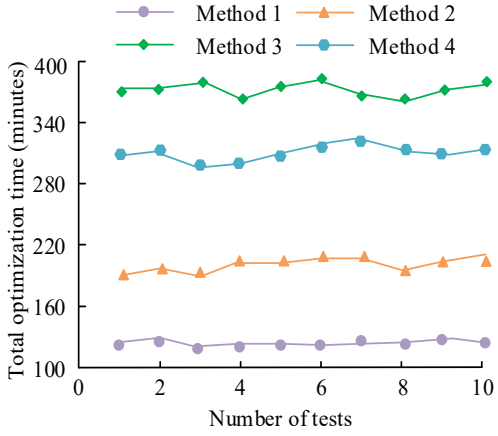
(b) Performance comparison of different methods under complex working conditions

Fig. 11. Performance comparison of different methods under standard and complex working conditions

The study ultimately compared the number of optimization iterations and total optimization time of each method in 10 independent tests, as shown in Figure 12. Figure 12(a) shows that Method 1 required the fewest iterations and was the most stable in finding the optimal hyperparameters, with an average of approximately 24 iterations. Methods 2, 3, and 4 had average iterations of 31, 68, and 91, respectively. Figure 12(b) shows that Method 1 completed the optimization in an average of approximately 125.1 minutes, the shortest time.



(a) Comparison of optimization iteration times for different methods



(b) Comparison of total optimization time using different methods

Fig. 12. Comparison of optimization efficiency of different FD methods

Method 2 had an average time of 199.5 minutes. In summary, the proposed BO-CapsNet method has an overwhelming advantage in computational efficiency. Its BO strategy can quickly find the optimal solution.

To further validate BO-CapsNet's adaptability in modern predictive maintenance scenarios, the study supplemented the optimization efficiency analysis presented in Figure 12 with robustness tests under domain shift, noise interference, and cross-machine generalization conditions, as shown in Table 2. Under standard operating conditions, BO-CapsNet achieved accuracy, F1 score, and recall rates of 99.7%,99.7%, and 99.8% respectively, with a loss value of only 0.10. When testing conditions expanded from uniform loads to load shifts and rotational speed variations, the model maintained accuracies of 97.6% and 96.8%, demonstrating that the Bayesian-optimized hyperparameter combination exhibits strong domain transfer stability without overfitting to specific conditions. Under noisy environments, model performance declined with decreasing signal-to-noise ratios but still achieved 94.3% accuracy and 94.0% F1 score under severe noise conditions, indicating that multi-scale convolutional networks effectively extract impact features across different sensory fields while channel attention mechanisms suppress noise-dominated redundant channels. In cross-machine testing, accuracy dropped to 92.6% yet retained high recall rates, highlighting how CapsNet's dynamic routing mechanism preserves hierarchical relationships among fault features and mitigates data distribution disparities across machine platforms. Overall, BO-CapsNet demonstrates consistent diagnostic performance under domain shift, noise interference, and cross-machine generalization conditions, confirming its engineering robustness for predictive maintenance applications.

4. DISCUSSION AND CONCLUSION

To solve the low accuracy of rolling BFD under complex operating conditions, this study designed a BO-CapsNet-based FD method for RBs. This method first constructed an improved CapsNet model integrating multi-scale convolution and CAMs. Then, the BO framework was introduced to automate hyperparameter optimization, forming a complete intelligent diagnosis process. Outcomes denoted that the improved CapsNet model outperformed CapsNet, SCNN, and DCNN models.

Table 2. Robustness verification of BO-CapsNet under different disturbance conditions

Testing condition	Domain shift setting	Noise level	Cross-machine setting	Accuracy/%	F1-score/%	Recall/%	Loss
Standard condition	Same load and same speed	0 dB	Same bearing platform	99.7	99.7	99.8	0.1
Load-shift condition	0 hp to 1 hp	0 dB	Same bearing platform	97.6	97.4	97.9	0.22
Speed-shift condition	1797 r/min to 1772 r/min	0 dB	Same bearing platform	96.8	96.5	97.1	0.29
Weak-noise condition	Same load and same speed	5 dB	Same bearing platform	96.1	95.9	96.4	0.34
Strong-noise condition	Same load and same speed	0 dB	Same bearing platform	94.3	94	94.7	0.47
Cross-machine condition	Different load and speed	5 dB	Different bearing platform	92.6	92.1	93.4	0.61

Under complex operating conditions, its diagnostic accuracy remained at 94.8%, with faster convergence. The loss value dropped below 0.1 after 40 iterations. In 10 repeated experiments, the proposed BO-CapsNet diagnostic method achieved an average accuracy of 99.7% with minimal fluctuation. In terms of computational efficiency, the average time required for optimization was only 125.1 minutes, lower than other comparative methods. In summary, the proposed BO-CapsNet FD method, through innovative model structure and intelligent parameter optimization, effectively solves the problem of insufficient feature extraction in conventional methods under complex operating conditions, providing a highly efficient new approach for intelligent FD of RBs.

However, the research methodology relies on sufficient labeled training data, which may be challenging to obtain in practical industrial settings. It should also be noted that the BO framework employed primarily facilitates offline hyperparameter optimization during model training, with the core objective of achieving optimal parameter configurations within limited evaluation iterations. While this approach significantly enhances model performance and reduces manual parameter tuning costs, the current framework is not specifically designed for online real-time optimization scenarios. Consequently, in actual deployment, the BO optimization process is typically completed during the offline training phase, while only the trained diagnostic model is retained during online operation, ensuring high inference efficiency and engineering feasibility for fault diagnosis tasks.

From an application perspective, the proposed BO-CapsNet method is suitable for predictive maintenance scenarios such as rotating machinery systems in smart factories, aerospace rotating components, and marine propulsion systems. In these applications, diagnostic models primarily handle state awareness and fault identification tasks, while control decisions are typically executed by independent control systems. Therefore, the research focuses on balancing diagnostic accuracy with computational overhead, enhancing industrial deployment efficiency through a combination of offline optimization and online inference. Future work could explore integrating fault diagnosis results with adaptive control strategies, intelligent decision-making systems, and digital twin platforms to establish an integrated framework for perception, diagnosis, and control, thereby improving the autonomous operation capabilities of complex industrial systems.

Source of funding: *The research is supported by: Gansu Province Natural Science Foundation; Research and Design of Condition Monitoring and Fault Diagnosis System for Heavy Scraper Conveyor, Project No.: 26JRRA854.*

Data availability: *The datasets used and/or analysed during the current study available from the corresponding author on reasonable request.*

Declaration of competing interest: *The author declares no conflict of interest.*

REFERENCES

1. Singh A, Madaan G, Hr S, Kumar A. Smart manufacturing systems: a futuristics roadmap towards application of industry 4.0 technologies. International Journal of Computer Integrated Manufacturing. 2023;36(3):411-428. https://doi.org/10.1080/0951192X.2022.2090607.
2. Konur S, Lan Y, Thakker D, Morkyani G, Polovina N, Sharp J. Towards design and implementation of Industry 4.0 for food manufacturing. Neural Computing and Applications. 2023;35(33):23753-23765. https://doi.org/10.1007/s00521-021-05726-z.
3. Liu C, Zheng P, Xu X. Digitalisation and servitisation of machine tools in the era of Industry 4.0: a review. International journal of production research. 2023; 61(12):4069-4101. https://doi.org/10.1080/00207543.2021.1969462.
4. Tao H, Qiu J, Chen Y, Stojanovic V, Cheng L. Unsupervised cross-domain rolling bearing fault diagnosis based on time-frequency information fusion. Journal of the Franklin Institute. 2023; 360(2):1454-1477. https://doi.org/10.1016/j.jfranklin.2022.11.004.
5. Keshun Y, Puzhou W, Yingkui G. Toward efficient and interpretative rolling bearing fault diagnosis via quadratic neural network with Bi-LSTM. IEEE Internet of Things Journal. 2024;11(13):23002-23019. https://doi.org/10.1109/JIOT.2024.3377731.
6. Wang X, Bian Q, Gao X, Zhao, C, Liu, M, Xie, X, Jiao, B. Study on the vibration characteristics of double-row tapered roller bearings for high-speed rail axle box. Industrial Lubrication and Tribology. 2024; 76(5):678-687. https://doi.org/10.1108/ILT-03-2024-0085.
7. Zhang Q, Deng L. An intelligent fault diagnosis method of rolling bearings based on short-time Fourier transform and convolutional neural network. Journal of Failure Analysis and Prevention. 2023;23(2): 795-811. https://doi.org/10.1007/s11668-023-01616-9.
8. Liu X, Xia L, Shi J, Zhang, L, Bai L, Wang S. A fault diagnosis method of rolling bearing based on improved recurrence plot and convolutional neural network. IEEE Sensors Journal. 2023;23(10): 10767-10775. https://doi.org/10.1109/JSEN.2023.3265409.
9. Zhao H, Wang L, Zhao Z, Deng W. A new fault diagnosis approach using parameterized time-reassigned multisynchrosqueezing transform for rolling bearings. IEEE Transactions on Reliability. 2024;74(1):2363-2372. https://doi.org/10.1109/TR.2024.3371520.
10. Zhao X, Wang L, Yang M, Chen Y, Xiang J. A novel small-sample fault diagnosis method for rolling bearings via continuous wavelet transform and Siamese neural network. IEEE Sensors Journal. 2024;24(15):24988-24996. https://doi.org/10.1109/JSEN.2024.3409768.
11. Zhang X, Kong J, Zhao Y, Zhao Y, Qian W, Xu X. A deep-learning model with improved capsule networks and LSTM filters for bearing fault diagnosis. Signal, Image and Video Processing. 2023;17(4):1325-1333. https://doi.org/10.1007/s11760-022-02340-x.

12. Shi X, Li T, Fang F, Zhu Y, Yang W, Luo B. Dissolved gas analysis for power transformer fault diagnosis combining domain knowledge and capsule network. IEEE Transactions on Dielectrics and Electrical Insulation. 2024;31(6):3386-3395. https://doi.org/10.1109/TDEI.2024.3374246.
13. Wang Y, Chen L. A multi-scale spatial–temporal capsule network based on sequence encoding for bearing fault diagnosis. Complex & Intelligent Systems. 2024;10(5):6189-6212. https://doi.org/10.1007/s40747-024-01462-8.
14. Chai J, Zhao X, Cao J. Small-sample fault diagnosis study of rolling bearings based on a residual parameterised convolutional capsule network. Insight-Non-Destructive Testing and Condition Monitoring. 2024;66(4):215-225. https://doi.org/10.1784/insi.2024.66.4.215.
15. Senthilnathan N, Babu TN, Varma KSD, Rushmith S, Reddy JA, Kavitha KVN, Prabha DR. Recent advancements in fault diagnosis of spherical roller bearing: A short review. Journal of Vibration Engineering & Technologies. 2024;12(4):6963-6977. https://doi.org/10.1007/s42417-024-01293-4.
16. Pu H, Zhang K, An Y. Restricted sparse networks for rolling bearing fault diagnosis. IEEE Transactions on Industrial Informatics. 2023;19(11):11139-11149. https://doi.org/10.1109/TII.2023.3243929.
17. Doudesis D, Lee KK, Boeddinghaus J, Bularga A, Ferry AV, Tuck C, Mills NL. Machine learning for diagnosis of myocardial infarction using cardiac troponin concentrations. Nature Medicine. 2023; 29(5):1201-1210. https://doi.org/10.1038/s41591-023-02325-4.
18. Diana Andrushia A, Mary Neebha T, Trephena Patricia A, et al. Image-based disease classification in grape leaves using convolutional capsule network. Soft Computing. 2023;27(3):1457-1470. https://doi.org/10.1007/s00500-022-07446-5.
19. Koy A, Çolak AB. The intraday high-frequency trading with different data ranges: A comparative study with artificial neural network and vector autoregressive models. Archives of Advanced Engineering Science. 2024;2(3):123-133. https://doi.org/10.47852/bonviewAAES32021325.
20. Gnanasambandam R, Shen B, Law ACC, Dou C, Kong Z. Deep Gaussian process for enhanced Bayesian optimization and its application in additive manufacturing. IISE Transactions. 2025;57(4):423-436. https://doi.org/10.1080/24725854.2024.2312905.
21. Boulkroune A, Boubellouta A, Bouzeriba A, Zouari F. Practical finite-time fuzzy synchronization of chaotic systems with non-integer orders: two chattering-free approaches. Journal of Systems Science and Systems Engineering. 2025;34(3):334-359. https://doi.org/10.1007/s11518-024-5635-7.
22. Rigatos G, Abbaszadeh M, Busawon K, Dala L, Zouari PF. Flatness-based control in successive loops for autonomous quadrotors. Journal of Dynamic Systems, Measurement, and Control. 2024,146(2):24501-24508. https://doi.org/10.1115/1.4063907.

Shanshan LI, Date of Birth: December 1981, Gender: Female, Place of Origin: Tianshui, Gansu Province, Ethnicity: Han. Master's degree in Mechanical Manufacturing and Automation from Taiyuan University of Technology in 2020, research focus: Mechanical Manufacturing and Automation.Work Experience: 2004 - Present, Teacher at Lanzhou Vocational Technical College.

Academic Background: Academic Publications: Published 2 papers in Peking University Chinese Core Journals, 2 papers in Chinese Science & Technology Core Journals, and more than 10 papers in provincial journals. Academic Textbooks: Chief editor of 1 textbook, deputy editor of 4 textbooks. Research Projects: Led 2 provincial-level projects, 3 municipal-level projects, and participated in more than 10 provincial and municipal projects.
e-mail: 18809319966@163.com