Research on Engineering Structures & Materials
www.jresm.org
Research Group

Research Article

Automobile gear fault detection based on VMD and ALO optimized SSDA network

Ling Huang [*,a]

Department of Mechanical and Electronic Engineering, Fushun Vocational Technology Institute (Fushun Teachers College), Fushun, China

Article Info

Article History:

Received 07 Aug 2026
Accepted 08 Oct 2026

Keywords:

Automobile gear;
Fault detection;
Variational mode decomposition;
Maximum correlation kurtosis deconvolution;
Stacked sparse denoising autoencoder network;
WOA

Abstract

Automobile gear transmission systems are vulnerable to wear, cracks, and broken teeth under long-term alternating loads, lubrication fluctuations, and assembly errors. Weak impact signals become difficult to detect due to noise interference and variable working conditions. This study proposes a fault detection method combining whale optimization algorithm, variational mode decomposition (VMD), maximum correlation kurtosis deconvolution (MCKD), and ant lion optimizer (ALO)-optimized stacked sparse denoising autoencoder (SSDA). For the representative tooth root crack condition at 1800 r/min and 50% rated load, the optimized model achieves an envelope kurtosis of 7.81, a main frequency energy ratio of 42.8%, and a decomposition signal-to-noise ratio of 14.63 dB, with clear fault impact peaks near sampling points 256, 768, 1280, and 1792. The detection model achieves accuracy, precision, recall, comprehensive evaluation index, and area under the curve values of 98.41%, 98.22%, 98.11%, 98.16%, and 99.40%, respectively. The inference time is only 6.38 ms, and stability reaches 97.36%. Under mixed variable operating conditions and -5 dB noise, recognition accuracy remains 97.81% and 93.48%, respectively. The proposed method effectively enhances weak fault feature extraction, cross-condition detection accuracy, and noise resistance for automobile transmission system monitoring.

© 2026 MIM Research Group. All rights reserved.

1. Introduction

Automobile gears are important basic components in the automobile transmission system to achieve power transmission, speed regulation and torque conversion. They are widely used in transmissions, differentials, drive axles and various transmission mechanisms. With the advancement of automobiles towards high speed, lightweight and high reliability, gears have been exposed to high-speed meshing, alternating loads, shock vibration and complex lubrication conditions for a long time, and their tooth surfaces are prone to wear, pitting, spalling, cracks and broken teeth. Once the fault is not identified in time, it will not only lead to reduced transmission efficiency, increased noise, and increased energy consumption, but may also cause instability in the transmission chain, chain damage to components, and even vehicle safety risks [1-2].

To address the challenge of gear fault diagnosis in high-noise environments, Goswami P et al. designed a method that combined minimum entropy deconvolution with adaptive noise fully integrated empirical mode decomposition. They used approximate entropy, dynamic time warping and fault correlation factors to select effective components. Research results showed that this method effectively retained gear health information in denoising, was superior to existing algorithms in wear and missing tooth fault detection, and performed robustly under different speed and load conditions [3]. To solve the problem that gear slack faults in traction motor drive systems

*Corresponding author: huangling0829@126.com
[a]orcid.org/ 0009-0003-9113-7837
DOI: http://dx.doi.org/10.17515/resm2026-2126vk0807rs

were easily misdiagnosed as wheelset idling or sensor faults, Ni Q et al. achieved a real-time fault detection employing the multi-dimensional information time series event model of the traction control unit. Results showed that this method effectively identified gear slack faults and had better performance in terms of sensitivity and reliability than existing methods [4]. To solve the difficulty in fault detection of cycloidal gears due to low speed and incomplete cycle compound motion under industrial robot working conditions, Shen Y et al. achieved a tooth fault detection relying on optimizing and improving envelope spectrum sideband energy ratio. The instantaneous angular velocity was estimated through the encoder signal and the signal was constructed to extract fault characteristics. Research results showed that this method effectively detected tooth profile faults of cycloid gears under robot operating conditions and was better than the cyclic spectral coherence method [5]. To sum up, although the above researches have achieved relatively good results, unresolved issues remain, such as insufficient generalization ability, high training cost and reliance on experience in parameter configuration when the sample distribution shifts, the speed load changes significantly, or in low signal-to-noise ratio scenarios.

To solve these shortcomings, the study constructs a fault feature extraction method grounded in Whale Optimization Algorithm (WOA), Variational Modal Decomposition (VMD) and Maximum Correlation Kurtosis Deconvolution (MCKD). The study further establishes the Ant Lion Optimizer (ALO) to optimize Stacked Sparse Denoising Autoencoder (SSDA) fault detection model. First, WOA serves to globally refine the key parameters of VMD to obtain a modal decomposition result that is more suitable for gear vibration signals. Then, weighted kurtosis is used to screen sensitive modes, and periodic impact characteristics are enhanced through MCKD. Finally, ALO is used to adaptively optimize the learning rate, number of hidden layer nodes, regularization coefficients and sparse constraint parameters of the SSDA network to achieve deep expression of fault characteristics and identification of cross-working conditions. The research aims to provide technical support for early fault detection of automobile gears, health monitoring of transmission systems and intelligent maintenance decisions. Research innovations are mainly reflected in three aspects: forming a parameter-adaptive joint mechanism of signal decomposition and impact enhancement, collaboratively modeling sensitive mode screening, periodic impact enhancement, and deep feature learning, and using ALO to improve the stability of the SSDA network under changing operating conditions and noise interference.

2. Methods

2.1 Automobile Gear Fault Detection Based on WOA-VMD-MCKD and ALO-Optimized SSDA

During the long-term meshing and transmission process, automobile gears are susceptible to factors such as alternating loads, fluctuations in lubrication status, manufacturing errors, and assembly deviations, leading to tooth surface pitting, wear, cracks, and broken teeth. Gear fault signals usually exhibit strong non-stationarity, weak impact components, strong background noise, and multi-modal coupling. It is difficult to fully characterize early fault information by directly using time domain or frequency domain features [6-7]. To achieve higher accuracy in fault feature extraction, VMD is employed to adaptively decompose original vibration signal and the complex signal into several intrinsic modal components with center frequencies, thereby reducing noise interference and highlighting fault impact characteristics in different frequency bands. The basic processing flow of VMD is shown in Figure 1.

As shown in Figure 1, the original vibration signal first enters the mode separation stage, and key parameters such as the number of decomposed modes, penalty factors, and initial center frequency are determined through parameter initialization, which provides a basis for subsequent operations. Then the bandwidth constraint minimization stage is entered, the signal decomposition problem is transformed into a variational optimization problem, and an augmented Lagrangian function is constructed to constrain the limited bandwidth characteristics of each modal component. The right side is the iterative optimization process. The system sequentially updates each modal component, central frequency and Lagrange multiplier under the alternating iterative update framework, and determines whether the stopping condition is met through the

convergence criterion. If it does not converge, it will continue to return to the iterative update. If the convergence conditions are met, the final decomposition result will be output.

Although VMD can decompose complex vibration signals into a set of intrinsic modal components with limited bandwidth, its decomposition effect has a strong dependence on the selection of mode number and penalty factor [8-9]. Therefore, to reduce the uncertainty caused by artificial experience parameter selection, the study introduces WOA to globally optimize the key parameters of VMD. By simulating the process of whales surrounding prey, spiral predation and random search, the optimal parameter is obtained under the guidance of fitness function, so that the decomposition results can more accurately retain the gear failure impact components and suppress noise interference. The parameter optimization process grounded in WOA-VMD is presented in Figure 2.

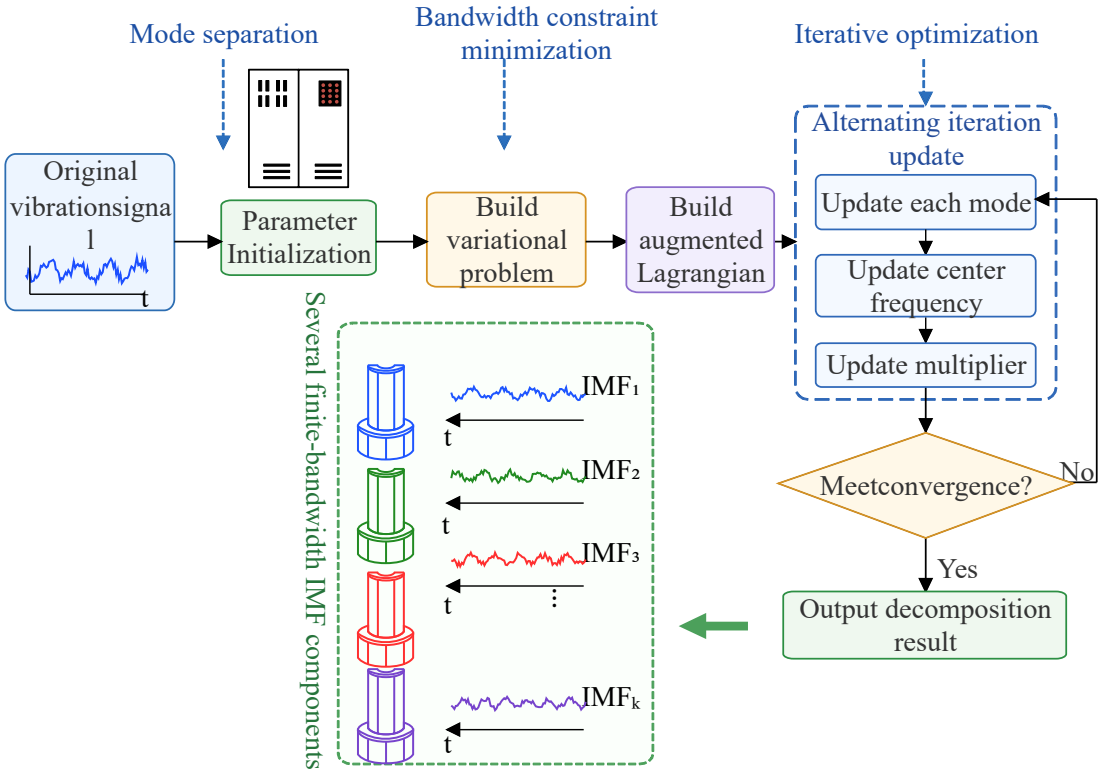


Fig. 1. VMD structure diagram

From Figure 2, the core idea is to use WOA to adaptively optimize key parameters in VMD to avoid relying on manual experience settings for mode numbers and penalty factors, which may lead to under-decomposition, over-decomposition, or weakening of fault impact components. First, the original vibration signal is input into the parameter encoding module. Each whale individual is defined as a set of VMD parameter combinations, as shown in Equation (1) [10-11].

$$X_i = (K_i, \alpha_i) \tag{1}$$

In Equation (1), X_i represents the i -th whale individual position. K_i denotes the VMD mode number, and α_i denotes the penalty factor. In the WOA search process, the mode number was constrained to $K_i \in [2,10]$, while the penalty factor was searched within $\alpha_i \in [100,5000]$. Since K_i is a discrete parameter, its updated value was rounded to the nearest integer before VMD decomposition. Subsequently, WOA evaluates signal decomposition quality under different parameter combinations according to the fitness function, and its fitness value is shown in Equation (2).

$$F_i = 0.40\widetilde{H_{e,\iota}} + 0.35\left(1 - \widetilde{K_{u,\iota}}\right) + 0.25\left(1 - \widetilde{R_\iota}\right) \tag{2}$$

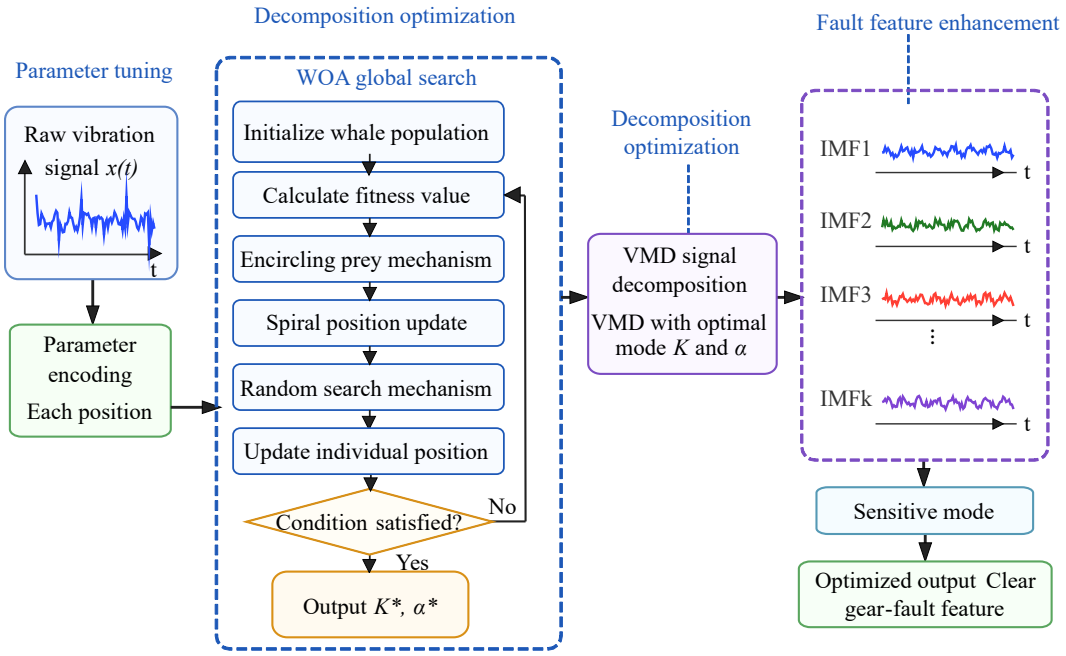


Fig. 2. Parameter optimization process based on WOA-VMD

In Equation (2), F_i denotes the fitness value of the i -th whale individual. $\widetilde{H_e}$ and R_i represent the normalized envelope entropy, kurtosis, and correlation coefficient, respectively. The corresponding weight coefficients are fixed at 0.40, 0.35, and 0.25 throughout the optimization process. Since lower envelope entropy and higher kurtosis and correlation indicate better decomposition quality, WOA searches for the parameter combination that minimizes F_i. In the global optimization stage, individual whales continuously adjust their positions by surrounding prey, spiral updates and random searches. The position update process is shown in Equation (3).

$$X(t+1) = X^*(t) - A \cdot |C \cdot X^*(t) - X(t)| \tag{3}$$

In Equation (3), $X(t)$ refers to the individual position at t iteration. $X(t+1)$ refers to the updated individual position. $X^*(t)$ refers to the current optimal individual position. A and C represent search control coefficients, which are used to adjust global exploration and local development capabilities. When the termination condition is reached, the algorithm outputs the optimal parameter combination and inputs it into the VMD module, whose expression is shown in Equation (4) [12].

$$(K^*, \alpha^*) = argmin\, f\,(K, \alpha) \tag{4}$$

In Equation (4), K^* represents the optimal mode number. α^* represents the optimal penalty factor. α^* represents the parameter combination that makes the fitness function obtain the optimal value. Finally, VMD uses optimal parameters to effectively decompose the original vibration signal and obtains multiple eigenmodal components, as presented in Equation (5).

$$x(t) = \sum_{k=1}^{K^*} u_k(t) + r(t) \tag{5}$$

In Equation (5), $x(t)$ represents the original vibration signal. $u_k(t)$ represents the kth eigenmodal component. K^* represents optimal modal components number. $r(t)$ represents the decomposition residual term. For consistency with Figures 1 and 3, the modal components denoted by $u_1(t), u_2(t), \ldots, u_k(t)$ in Equation (5) correspond respectively to IMF1, IMF2, ..., IMFk in the flowcharts.

WOA-VMD can improve the modal decomposition effect of vibration signals through adaptive parameter optimization, so that the original gear fault signal is decomposed into multiple eigenmodal components with different center frequencies. However, under actual automobile gear operating conditions, the impact signals generated by early pitting corrosion, cracks and slight wear are often of low amplitude, and are easily interfered by background noise of the transmission system, speed fluctuations and structural resonance components. Relying only on optimal decomposition may still cause problems such as insufficient prominent fault impact components, insufficiently stable sensitive mode selection, and insufficient expression of periodic features. Therefore, after obtaining the WOA-VMD decomposition results, MCKD is further introduced to perform impact enhancement and noise suppression on sensitive modes containing fault information, as shown in Figure 3.

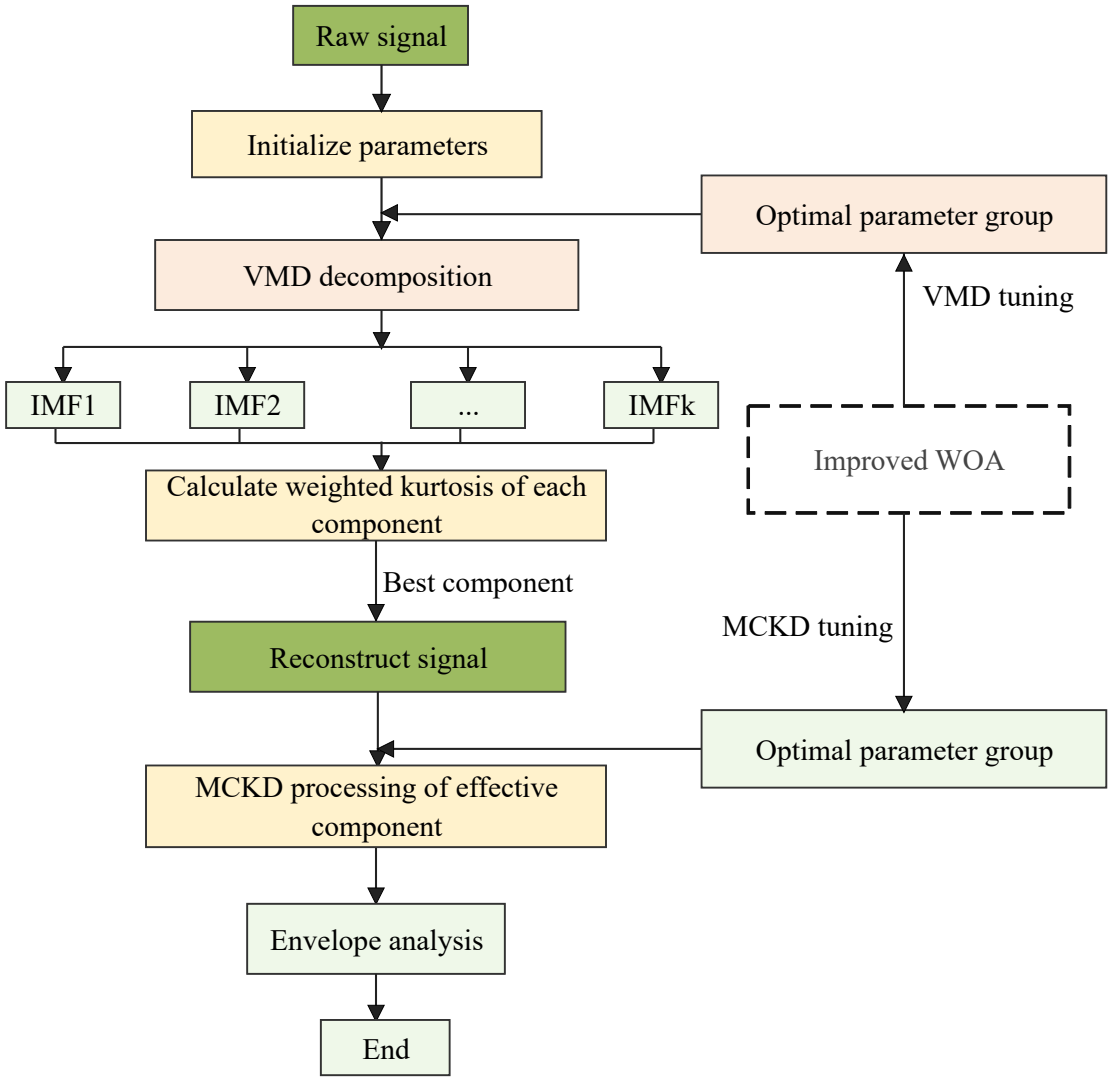


Fig. 3. Automobile gear fault feature extraction process based on WOA-VMD-MCKD

From Figure 3, the original vibration signal first enters the parameter initialization stage, and the optimal parameter combination of VMD is determined under the action of the WOA optimization mechanism, including the mode number and penalty factor. Therefore, it reduces the impact of artificial parameter setting on the decomposition results. Subsequently, VMD decomposes the signal to obtain multiple intrinsic mode components, such as IMF1, IMF2 to IMFk. Different components correspond to different frequency band information in the original signal. To screen out sensitive components that are more closely related to gear faults, the weighted kurtosis of each component is calculated by jointly considering the kurtosis of the modal component and its correlation with the original vibration signal. Specifically, the kurtosis is used to characterize the impulsiveness of each mode, while the normalized correlation coefficient is introduced as the weighting factor to evaluate the preservation of fault-related information. A larger weighted

kurtosis therefore indicates that the corresponding mode contains stronger impulsive characteristics while maintaining a higher correlation with the original signal, and the component with the highest weighted kurtosis is selected for signal reconstruction. The reconstructed signal may still contain background noise and non-fault impacts, so MCKD is continued to be introduced for effective modal component processing. Meanwhile, the MCKD parameters are optimized through WOA to obtain the best parameter combination and improve the enhancement effect of periodic impact components. Finally, envelope analysis is conducted on the enhanced signal to extract the gear fault characteristic frequency and its related modulation information.

2.2 Automobile Gear Fault Detection Model Based on ALO-SSDA

Through the WOA-VMD-MCKD feature extraction process, the noise interference in the original automobile gear vibration signal is suppressed, and the weak periodic impact components and sensitive fault features are enhanced. This provides more discriminative input information for subsequent fault identification [13-14]. However, in the actual automobile gear operating environment, different speeds, loads, and collection conditions will lead to distribution differences between training data and test data. If the traditional supervised learning model is directly used for classification, it is easy to have the problem that the model fully fits the source domain samples but has insufficient generalization ability for the target domain samples. To improve stability of the model under varying operating conditions, research further builds an automobile gear fault detection model based on SSDA. Its structure is shown in Figure 4.

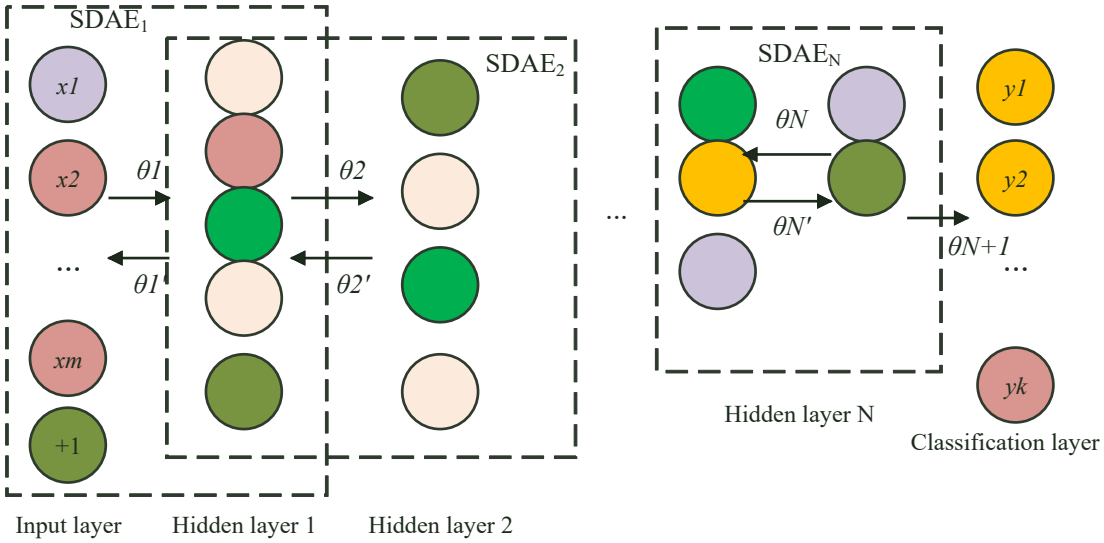
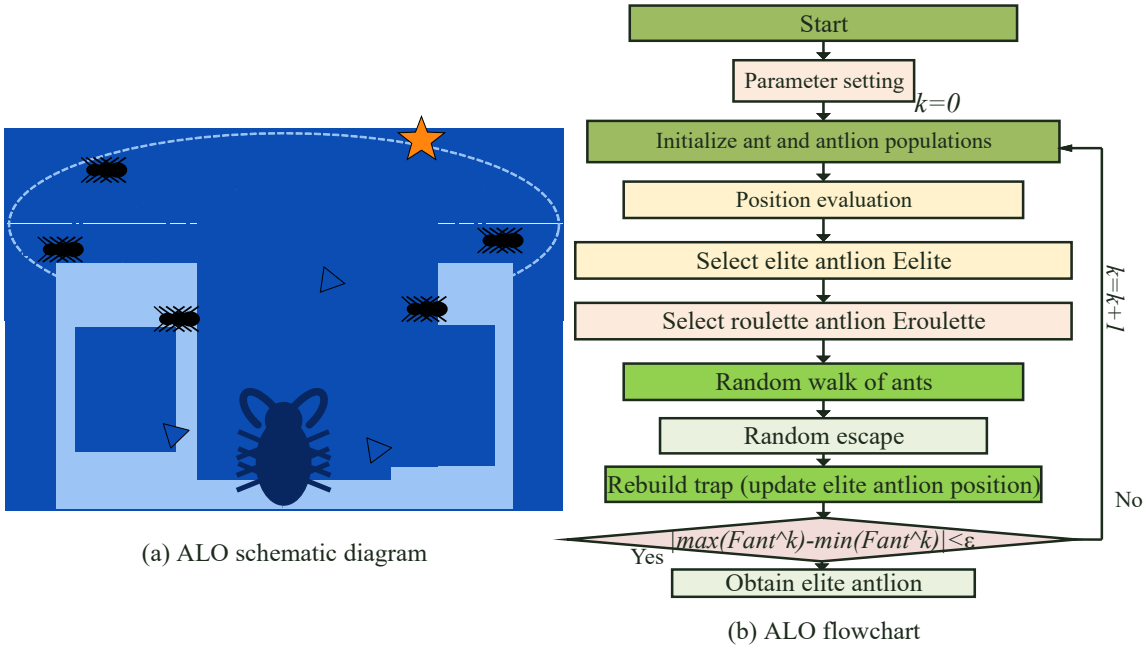


Fig. 4. Automobile gear fault detection model based on SSDA

From Figure 4, the model includes an input layer, multiple stacked sparse autoencoder hidden layers, and a classification layer. The input layer receives the gear fault feature vector processed by WOA-VMD-MCKD, including multiple time-frequency features, impact features, and reconstructed signal features. Subsequently, the features enter the first hidden layer, the second hidden layer, and the N-th hidden layer in sequence. The number of stacked hidden layers N is fixed during model construction, while all hidden layers adopt the same number of neurons. Each sparse autoencoder layer performs feature compression, reconstruction, and nonlinear mapping through the encoding and decoding process. The forward parameters are used to extract high-level abstract features, and the reverse parameters are used to constrain the reconstruction error, allowing the network to retain key fault information while suppressing redundant noise. As the number of network layers deepens, the original features are gradually transformed from low-level local descriptions into high-level semantic expressions, improving the separability between different gear fault types. Finally, the classification layer outputs corresponding fault categories based on deep features, such as normal, worn, cracked, and broken teeth.

Although SSDA can achieve deep expression and classification identification of automobile gear fault characteristics through a multi-layer sparse autoencoding structure, its detection performance is greatly affected by network hyperparameters. If the number of hidden layer nodes,

learning rate, sparse constraint coefficient, regularization parameters and classifier parameters are set unreasonably, it will easily lead to slow convergence of the model, insufficient feature expression or over-fitting, thus weakening the stability of fault identification under variable working conditions [15-16]. To further improve the parameter adaptive capability of the SSDA network, the study introduces ALO to globally optimize key hyperparameters.



(a) ALO schematic diagram

(b) ALO flowchart

Fig. 5. ALO optimization module process

From Figure 5, parameters such as the number of ants and ant lions, the maximum number of iterations, random walk function, the escape probability and the convergence threshold are first set. Iterations number is set to 0. Then ant and ant lion populations are initialized, where each individual represents a set of SSDA parameter combinations to be optimized, as shown in Equation (6).

$$X_i^k = [h_i^k, \eta_i^k, \lambda_i^k, \rho_i^k] \tag{6}$$

In Equation (6), X_i^k refers to position of the i -th individual in the k -th iteration. h_i^k refers to hidden layer nodes number. η_i^k represents the learning rate. λ_i^k represents the regularization coefficient, and ρ_i^k represents the sparse constraint coefficient. After completing the initialization, the algorithm evaluates the positions of ants and ant lions, and the fitness is given by Equation (7) [17].

$$F_i^k = \phi(X_i^k) \tag{7}$$

In Equation (7), F_i^k refers to fitness value of the i -th individual at the k -th iteration. $\phi(\cdot)$represents evaluation function composed of the verification set error, classification accuracy or loss function. Then, natural and roulette elite ant lions are selected according to their fitness, and the ants are guided to perform random walks and position updates, as shown in Equation (8).

$$M_i^{k+1} = \frac{R_A^k + R_E^k}{2} \tag{8}$$

In Equation (8), M_i^{k+1} refers to updated position of the i ant. R_A^k represents random walk position generated around the roulette ant lion. R_E^k represents the random walk position generated around the elite ant lion. If the fitness of ant is better than that of corresponding ant lion, random escape and trap reconstruction are triggered, and the ant lion's position is updated. The algorithm determines whether it has converged by the fitness difference, as shown in Equation (9).

$$max(F_{ant}^{k}) - min(F_{ant}^{k}) < \varepsilon \tag{9}$$

In Equation (9), F_{ant}^{k} represents the fitness of ant population in the k iteration, and F_{ant}^{k} represents the convergence threshold. When the conditions are met, the natural elite ant lion is output, whose position is the optimal parameter combination of SSDA. Otherwise, $k = k + 1$ is allowed to continue iterative optimization. During SSDA training, reconstruction learning, sparse feature constraint, fault classification, and subdomain distribution alignment are jointly optimized. The total loss function is defined as Equation (10)

$$L_{total} = L_{rec} + \alpha L_{sparse} + \beta L_{cls} + \gamma L_{MMD} \tag{10}$$

In Equation (10), when L_{rec} denotes the reconstruction loss of the stacked autoencoder, L_{sparse} represents the sparsity penalty used to constrain hidden-unit activation, L_{cls} denotes the fault classification loss, and L_{MMD} represents the Maximum Mean Discrepancy between source-domain and target-domain feature distributions. Minimizing the joint objective enables the network to retain fault-sensitive features while reducing the distribution discrepancy between different operating conditions.

In summary, the structure of the monitoring model established based on ALO-SSDA is shown in Figure 6.

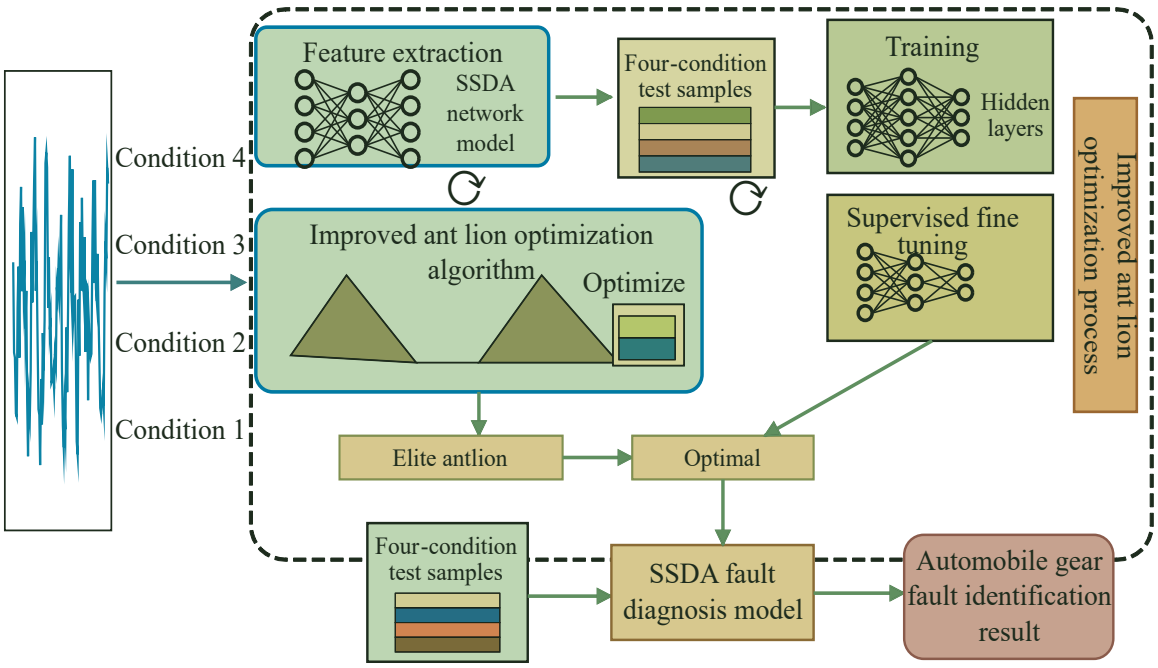


Fig. 6. Structure of automobile gear fault detection model based on ALO-SSDA

From Figure 6, the original gear vibration signal first enters the feature processing process through the input end. After the feature extraction is completed, the signal is input into the SSDA network model, and the deep expression and dimensionality reduction mapping of fault features are achieved through a multi-layer coding structure. Subsequently, the model uses hidden layer features to complete the discrimination of fault categories, and combines the reconstruction error and classification error to jointly constrain the network parameters during the training process. So the extracted features can not only retain the effective information in the original signal, but also have strong category discrimination ability. To avoid relying on manual experience settings for SSDA network parameters, the ALO module is further introduced to perform adaptive optimization of network weights, learning rates, number of hidden layer nodes and related control parameters. During the optimization process, fitness evaluation is used to screen better parameter

combinations, and the optimal parameters are fed back to the SSDA fault diagnosis model. Finally, the model outputs gear fault identification results based on the optimized deep features.

3. Results

3.1 Automobile Gear Fault Feature Extraction Method Based on WOA-VMD-MCKD

To validate the effectiveness of the automobile gear fault detection method based on VMD and ALO-optimized SSDA, an automobile gear fault test rig was established. The tested gearbox employed a spur-gear transmission pair with 32 teeth, a module of 2 mm, and a pitch circle diameter of 64 mm. The drive system was powered by a 3-kW variable-frequency motor, and the shaft speeds were set to 1200, 1800, and 2400 r/min. Four load torques of 10, 20, 30, and 40 N·m were applied through an electromagnetic loading device. Gearbox vibration signals were measured using a PCB Piezotronics 352C33 piezoelectric accelerometer with a nominal sensitivity of 100 mV/g, mounted radially on the bearing housing adjacent to the gearbox output shaft. Signal acquisition was performed using an NI PXIe-4492 dynamic signal acquisition module with a sampling frequency of 27.52 kHz. The experimental objects included four states: normal, tooth surface wear, tooth root crack, and broken tooth. A total of 3,000 samples were collected for each state, resulting in 12,000 samples, with each sample containing 2048 sampling points. The collection working conditions were set to three shaft speeds of 1200, 1800, and 2400 r/min and four load combinations.

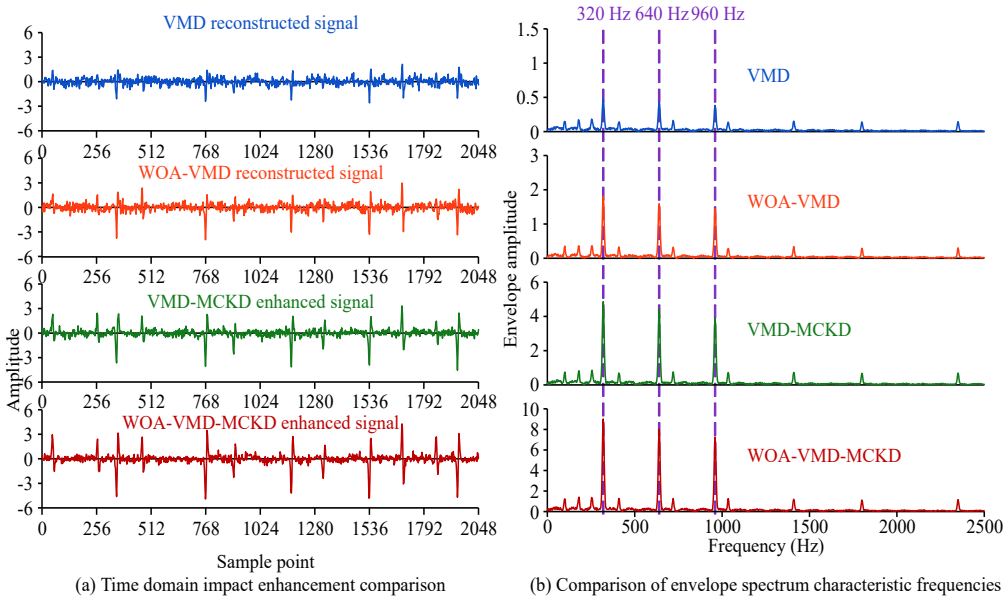


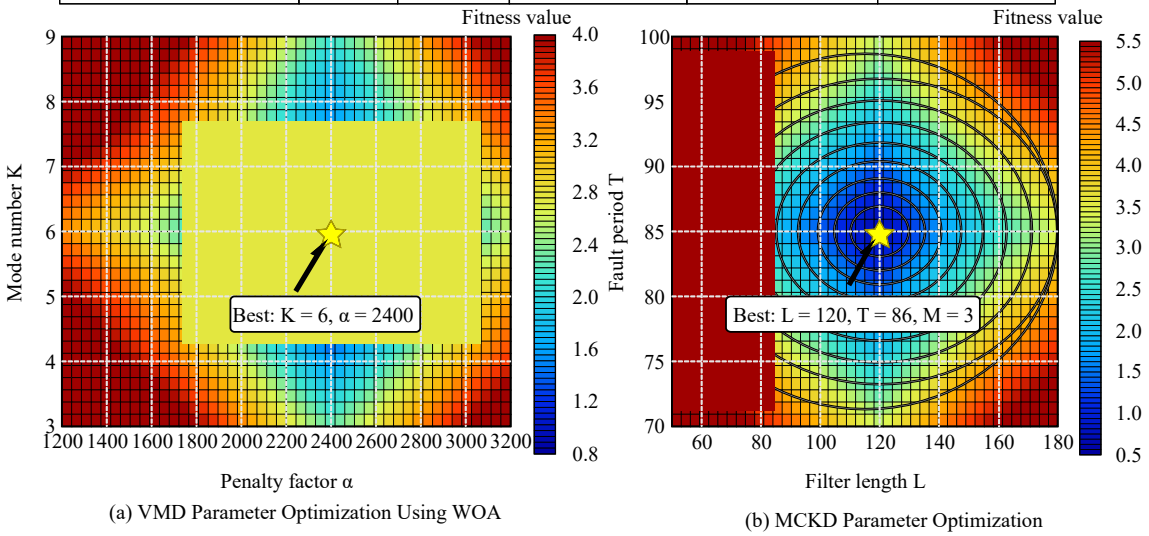
Fig. 7. Enhancement effects comparison of automobile gear fault signals under different feature extraction methods

The vibration signals were acquired at a sampling frequency of 27.52 kHz. A total of 3,000 samples were collected for each fault state, resulting in 12,000 samples, with each sample containing 2048 sampling points. The training set, validation set, and test set were split in a 6:2:2 ratio, and all experiments were independently repeated five times. The mean recognition accuracy and its standard deviation were calculated across the five runs. Stability was evaluated from the relative fluctuation of recognition accuracy among the five repeated experiments, with a higher Stability value indicating a smaller variation and more consistent model performance. A tooth root crack sample collected at 1800 r/min and 50% rated load was selected as the representative case for the feature enhancement analysis in Figures 7 and 8. VMD, WOA-VMD, VMD-MCKD and WOA-VMD-MCKD methods were selected as comparison models. The tested gearbox employed a spur-gear transmission pair, and the monitored gear had 16 teeth, a module of 2 mm, and a pitch circle diameter of 32 mm. The shaft speeds were set to 1200, 1800, and 2400 r/min, and the representative signal used for the frequency-domain analysis in Figures 7 and 8 was collected at

1200 r/min. The enhancement effects of automobile gear fault signals under different feature extraction methods are shown in Figure 7.

Figure 7(a) shows the time domain impact enhancement results, and Figure 7(b) shows the envelope spectrum characteristic frequency comparison results. From Figure 7(a), the VMD reconstructed signal still had many random fluctuations and the impact components were not concentrated enough. After WOA-VMD processing, the periodic impact began to increase, but some peaks were still interfered by noise. VMD-MCKD further improved the impact peak clarity. The periodic pulse of WOA-VMD-MCKD was the most obvious, with strong impact peaks appearing near the sampling points of about 256, 768, 1280 and 1792, indicating that this method better extracted the repetitive impact caused by local damage to the gear. From Figure 7(b), the four methods exhibited characteristic peaks near 320 Hz, 640 Hz, and 960 Hz. Under the representative speed of 1200 r/min, the rotational frequency of the monitored gear was 20 Hz. Considering the 16 teeth of the monitored gear, the theoretical gear mesh frequency was 320 Hz. Therefore, the spectral peaks at 320 Hz, 640 Hz, and 960 Hz were identified as the gear mesh frequency and its second and third harmonics, respectively. The VMD spectral peaks were relatively weak and accompanied by several spurious spectral lines, whereas the characteristic components became progressively clearer after WOA-VMD and MCKD processing. The main frequency peaks of WOA-VMD and VMD-MCKD gradually increased. WOA-VMD-MCKD had the highest peaks at the three characteristic frequencies, and the background spectral lines were lower. To sum up, WOA optimization improved the matching of VMD parameters, and MCKD enhanced the periodic impact component. The combination of the two could more effectively highlight the characteristics of automobile gear faults. The heat map was used for analysis, and the results are shown in Figure 8.

Method	K	α	Envelope kurtosis	Main frequency energy ratio	Decomposition SNR
VMD	5	2000	4.86	27.3%	9.74 dB
WOA-VMD	6	2400	5.92	34.6%	11.82 dB
VMD-MCKD	5	2000	6.31	37.9%	12.58 dB
WOA-VMD-MCKD	6	2400	7.81	42.8%	14.63 dB



(a) VMD Parameter Optimization Using WOA

(b) MCKD Parameter Optimization

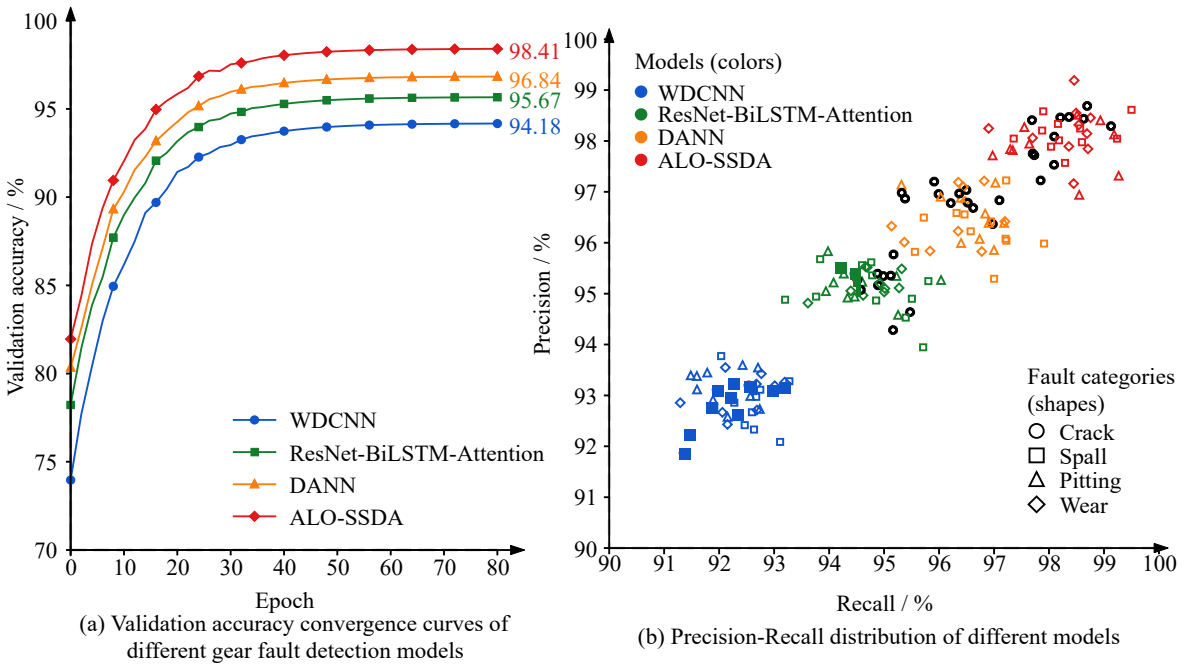
Fig. 8. WOA-VMD-MCKD parameter optimization and fitness distribution analysis

Figure 8(a) shows the optimization results of WOA on VMD mode number K and penalty factor α. Figure 8(b) shows the parameter optimization results of MCKD filter length L and fault period T. From Figure 8(a), the fitness value reached the lowest area near K=6 and α=2400. When K was too small or α was set unreasonably, the fitness value increased, and modal aliasing or insufficient bandwidth constraints were prone to occur. The envelope kurtosis of WOA-VMD-MCKD reached 7.81, the main frequency energy ratio reached 42.8%, and the decomposed signal-to-noise ratio reached 14.63 dB, which were better than VMD, WOA-VMD and VMD-MCKD. From Figure 8(b), the MCKD parameters had the lowest fitness when L=120, T=86, and M=3. For the representative tooth root crack signal at 1800 r/min, the fundamental fault characteristic frequency was 320 Hz. In

summary, WOA optimization avoided the uncertainty caused by artificial parameter setting, made the VMD decomposition and MCKD impact enhancement processes more stable, and provided a better parameter basis for automobile gear fault feature extraction.

3.2 Performance Analysis of Automobile Gear Fault Detection Model Based on ALO-SSDA

To validate the classification of ALO-SSDA model in automobile gear fault detection, the study selected the Wavelet Decomposition Convolutional Neural Network (WDCNN), Residual Network - Bidirectional Long Short-Term Memory - Attention Mechanism (ResNet-BiLSTM-Attention) and Domain-Adversarial Neural Network (DANN) as a comparison model. The proposed ALO-SSDA was as the original model for comprehensive comparison.



(a) Validation accuracy convergence curves of different gear fault detection models

(b) Precision-Recall distribution of different models

Fig. 9. Performance comparison of different automobile gear fault detection models

Figure 9(a) shows the verification accuracy convergence curves between different models, and Figure 9(b) shows the distribution in the precision and recall dimensions. From Figure 9(a), the verification accuracy of the four models gradually converged with the increase of training rounds. Among them, the final accuracy of WDCNN was 94.18%, ResNet-BiLSTM-Attention was improved to 95.67%, and DANN reached 96.84%. ALO-SSDA finally reached 98.41%, and the curve became stable after about 40 rounds, indicating that the ALO-optimized SSDA network had better parameter configuration and convergence quality. From Figure 9(b), the WDCNN scatter points were mainly distributed in the recall rate of about 92% and the precision rate of about 93%, indicating that its ability to distinguish complex fault categories was limited. the DANN scatter points were concentrated in the recall rate of 96% to 97% and the precision rate of 96% to 97%, indicating that the domain adaptation mechanism improved cross-working condition recognition capabilities. ALO-SSDA scatter points were concentrated in the upper right corner, and the recall rate and precision rate were mostly maintained at about 98%. To sum up, ALO-SSDA was superior to other models in terms of convergence accuracy, category discrimination ability and test stability, and was more suitable for automobile gear fault detection tasks. The model performance was further analyzed, and the results are shown in Figure 10.

Figure 10(a) shows the accuracy comparison across working conditions, and Figure 10(b) shows the change in recognition accuracy under different signal-to-noise ratio levels. From Figure 10(a), under the same speed and load conditions, the accuracy of each model was high, among which ALO-SSDA reached 98.41% and DANN was 96.84%. When the working conditions migrated, the model performance declined, but A LO-SSDA still maintained 96.93%, 96.72% and 97.81% under low load to high load, high load to low load and mixed variable operating conditions, respectively, which was

higher than DANN's 94.58%, 94.21% and 95.17%. From Figure 10(b), as the signal-to-noise ratio decreased from 10 dB to -5 dB, the accuracy of each model continued to decrease, among which WDCNN decreased from 93.76% to 85.37%, which was a larger decline. ALO-SSDA still reached 93.48% at -5 dB. To sum up, ALO-SSDA showed better detection stability and engineering applicability in complex working conditions and low signal-to-noise ratio environments.

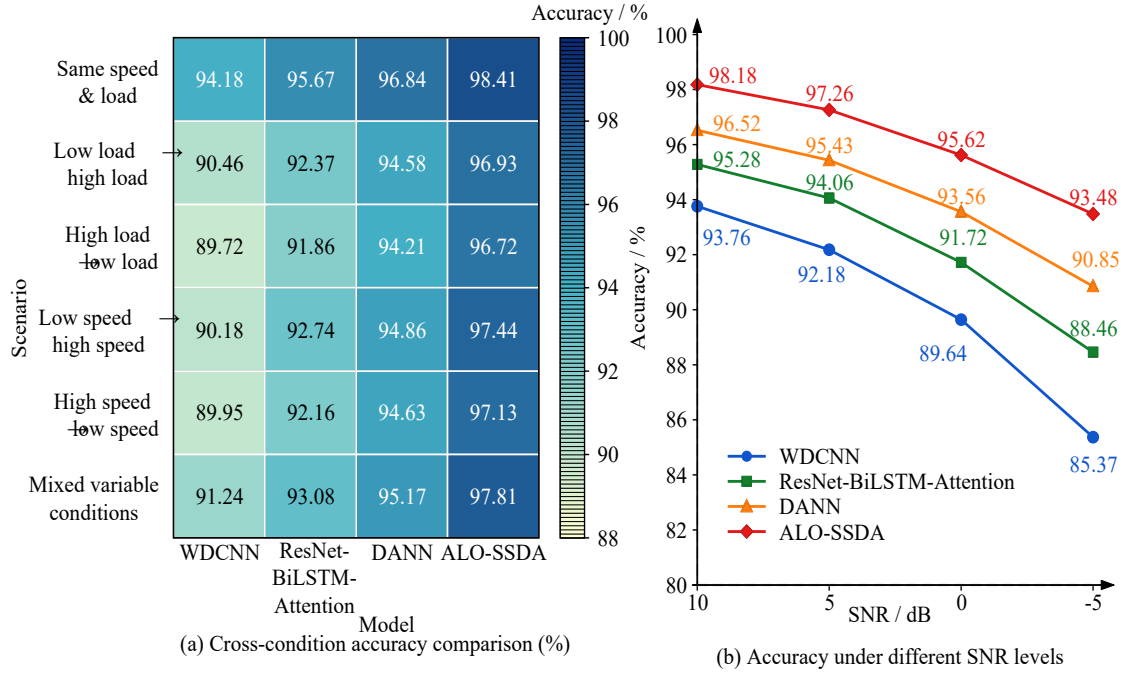


Fig. 10. Generalization performance analysis of different models under changing working conditions and noise interference conditions

The comprehensive performance was analyzed, as shown in Table 1.

Table 1. Comparison of comprehensive performance of different gear fault detection models

Metric	WDCNN	ResNet-BiLSTM-Attention	DANN	SSDA	PSO-SSDA
Accuracy/%	94.18	95.67	96.84	94.96	97.26
Precision/%	93.76	95.21	96.52	94.91	97.08
Recall/%	93.41	94.96	96.17	94.73	96.86
F1/%	93.58	95.08	96.34	94.82	96.97
AUC/%	96.8	97.6	98.5	97.31	98.82
Training time/s	186.4	248.7	276.3	218.5	238.6
Inference time/ms	5.82	7.46	8.24	6.12	6.31
Params/M	1.36	2.18	2.64	1.84	1.9
Stability/%	91.24	93.15	94.82	91.36	95.68

From Table 1, the standard SSDA achieved an accuracy of 94.96% and an F1 of 94.82%, while PSO-SSDA improved these values to 97.26% and 96.97%, respectively. ALO-SSDA further achieved an accuracy of 98.41%, precision of 98.22%, recall of 98.11%, F1 of 98.16%, and AUC of 99.40%. Compared with PSO-SSDA, ALO-SSDA improved the accuracy and F1 by 1.15 and 1.19 percentage points, respectively, while Stability increased from 95.68% to 97.36%. These results further distinguish the performance gain obtained by ALO-based hyperparameter optimization from that of conventional grid search and PSO optimization. Although it had certain local impact feature extraction capabilities, it was not adaptable to the time series correlation of complex non-stationary signals and the distribution changes across working conditions. In contrast, ALO-SSDA achieved an AUC of 99.40%, an inference time of only 6.38 ms, and a Stability of 97.36%, which was calculated from the variation in recognition accuracy across five independent repeated experiments, indicating that it had achieved a better balance between recognition accuracy, generalization ability and operating efficiency, and was more suitable for automobile gear fault

detection scenarios. For ALO-SSDA, the parameter count was calculated using the 2048-592-592-592-4 encoder-classifier architecture. The resulting number of trainable parameters in the fault-identification path was approximately 1.917 million, which was rounded to 1.92 M in Table 1. Ablation experiments analyzed each module of the proposed model, as shown in Table 2.

Table 2. Ablation experiment table

Model variant	SSDA	VMD-SSDA	WOA-VMD-SSDA	WOA-VMD-MCKD-SSDA	WOA-VMD-MCKD-ALO	WOA-VMD-MCKD-ALO-SSDA
WOA-VMD	×	√	√	√	√	√
MCKD	×	×	×	√	√	√
ALO	×	×	×	×	√	√
Subdomain adaptation	√	√	√	√	×	√
Feature energy ratio/%	31.26	35.47	38.92	42.8	42.8	42.8
Maximum Mean Discrepancy	0.184	0.161	0.139	0.118	0.127	0.086
Convergence epochs	67	61	56	52	48	43
Cross-condition accuracy/%	93.58	94.62	95.38	96.21	96.76	97.81
Noise accuracy at -5 dB/%	88.74	90.15	91.46	92.37	92.86	93.48
F1/%	94.82	95.43	96.12	96.95	97.31	98.16
Stability/%	91.36	92.54	93.87	95.08	95.74	97.36

From Table 2, after each module was gradually introduced, the model had been improved in terms of feature expression, domain adaptation, and complex working condition recognition. The Feature energy ratio of SSDA was only 31.26%, the cross-condition accuracy was 93.58%, and the Noise accuracy at -5 dB was 88.74%, which showed that it was difficult to fully extract the gear impact characteristics under low signal-to-noise ratio by relying on SSDA alone. After introducing VMD, the proportion of feature energy increased to 35.47%, indicating that modal decomposition enhanced effective frequency band information. After further using WOA-VMD, this index increased to 38.92%, and the Maximum Mean Discrepancy dropped to 0.139, indicating that parameter optimization reduced modal aliasing and improved inter-domain feature distribution. After adding MCKD, the Feature energy ratio reached 42.80% and F1 increased to 96.95%. Compared with WOA-VMD-MCKD-ALO-SSDA, after adding subdomain adaptation to the complete model, the MMD distance dropped from 0.127 to 0.086, the cross-condition accuracy increased to 97.81%, and the Stability reached 97.36%. To sum up, WOA-VMD, MCKD, ALO and SSDA subdomain adaptation modules all contributed to performance improvement, among which the complete model performed best in terms of variable working condition recognition, low-noise robustness and stability.

4. Conclusion

Aiming at the problems of weak impact of early failure of automobile gears, strong background noise, severe coupling of modal components, and inconsistent distribution of training samples and test samples under variable working conditions, this study achieves an automobile gear fault detection method on the basis of WOA-VMD-MCKD feature enhancement and ALO-SSDA fault identification. Experimental results show that WOA-VMD-MCKD forms clear impact peaks near sampling points such as 256, 768, 1280 and 1792, and the fault characteristic frequencies and frequency doubling components at 320 Hz, 640 Hz and 960 Hz in the envelope spectrum are more prominent. Parameter optimization results show that VMD achieves better decomposition effect near K=6, α=2400, and MCKD has the lowest fitness under the conditions of L=120, T=86, and M=3. The envelope kurtosis of WOA-VMD-MCKD reaches 7.81, the main frequency energy ratio reaches 42.8%, and the decomposed signal-to-noise ratio reaches 14.63 dB. In terms of fault detection performance, ALO-SSDA's Accuracy, Precision, Recall, F1 and AUC reach 98.41%, 98.22%, 98.11%, 98.16% and 99.40% respectively, which are 1.57, 1.70, 1.94, 1.82 and 0.90 percentage points higher

than DANN respectively. The inference time is only 6.38 ms and the stability reached 97.36%. In summary, the proposed method effectively solves the problems of difficulty in extracting weak features of gear faults and unstable identification of changing working conditions, and achieve a good balance between detection accuracy, convergence efficiency, noise immunity and engineering applicability. However, there are still deficiencies in the research. Experimental data mainly come from limited fault types and working condition combinations constructed on experimental benches, and have not fully covered complex scenarios such as long-term service wear, compound faults, extreme temperature rises and real road load impacts. Follow-up research can further expand the scale of real vehicle operating data and combine it with an online incremental learning mechanism to improve continuous adaptability to unknown faults and dynamic operating conditions.

References

[1] Hashim S, Mishra SK, Shakya P. An approach for adaptive filtering with reinforcement learning for multi-sensor fusion in condition monitoring of gearboxes. Computers in Industry. 2025;164(12):54-77. https://doi.org/10.1016/j.compind.2024.104214

[2] Li K, Qiao B, Wang Y. Non-convex group sparse regularization method for local fault detection of spiral bevel gear. Measurement. 2025;247(13):1168-1179. https://doi.org/10.1016/j.measurement.2025.116808

[3] Goswami P, Rai RN. A Novel Methodology for Chipped Teeth and Worn Tooth Gear Fault Detection Using Minimum Entropy Deconvolution and CEEMDAN. Journal of Vibration Engineering & Technologies. 2025;13(1). https://doi.org/10.1007/s42417-024-01622-7

[4] Ni Q, Liu J, Zhang J. A Traction Control System Signal-based Method for Gear Slack Fault Detection in Electric Locomotive. IEEE Transactions on Power Electronics. 2025;40(6):8766-8775. https://doi.org/10.1109/tpel.2025.3538789

[5] Shen Y, Guo Y, Fan J. Fault detection of RV reducer cycloidal gear under robot working conditions based on optimization improved envelope spectrum sideband energy ratio. Measurement Science and Technology. 2025;36(1):16155-16165. https://doi.org/10.1088/1361-6501/ad9102

[6] Zaiem K, Bouras H, Rachedi MF. Bayesian convolution neural network for mechanical transmission fault diagnosis with empirical signal processing. Journal of Mechanical Science and Technology. 2025;39(6):3131-3145. https://doi.org/10.1007/s12206-025-0512-y

[7] Bessaoudi M, Habbouche H, Benkedjouh T, Mesloub A. A hybrid approach for gearbox fault diagnosis based on deep learning techniques. The International Journal of Advanced Manufacturing Technology. 2024;133(5-6):2861-2874. https://doi.org/10.1007/s00170-024-13873-5

[8] Mrabti A, Younes R, Ouelaa N. Robust fault detection and severity classification in rotating machinery using VMD-LSTM for limited data scenarios. Advances in Mechanical Engineering. 2025;17(5). https://doi.org/10.1177/16878132251342909

[9] Yu X, Zhang Z, Tang BZM. A multi-head self-attention autoencoder network for fault detection of wind turbine gearboxes under random loads. Measurement Science & Technology. 2024;35(8):8613701-8613711. https://doi.org/10.1088/1361-6501/ad4dd4

[10] Zhang Y, Wang X, Wu Z. Simulation-driven fault detection for the gear transmission system in major equipment. Measurement and Control. 2024;57(9):1268-1285. https://doi.org/10.1177/00202940241230275

[11] Yang N, Liu Y, Wei Y. Iterative Laplace of Gaussian filter and improved Teager energy operator for bearing fault detection in gearboxes. Measurement Science & Technology. 2024;35(9):9613501-9613512. https://doi.org/10.1088/1361-6501/ad5903

[12] Shen Y, Guo Y, Fan J. Robot Joint Cycloidal Gear Tooth Fault Detection in Robot Working Conditions Based on Piecewise Mean Difference. IEEE Sensors Journal. 2024;24(20):32740-32747. https://doi.org/10.1109/JSEN.2024.3450862

[13] He X, Zhao F, Song N, Su C, Liu P. A deep-transfer-learning fault diagnosis method for gearboxes based on discriminative feature extraction and improved domain adversarial neural networks. Nondestructive Testing and Evaluation. 2026;41(3):1722-1743. https://doi.org/10.1080/10589759.2025.2495802

[14] Kalala GN, Boujelben A, Rasolofondraibe L. Towards Accurate Gear Fault Diagnosis: Hybrid Data Enrichment by Digital Twins for Machine Learning-Based Monitoring. Journal of Vibration Engineering & Technologies. 2026;14(1):13-14. https://doi.org/10.1007/s42417-025-02252-3

[15] Goswami P, Rai RN. Data-driven sensor selection for industrial gearbox fault diagnosis using principal component analysis. Measurement Science and Technology. 2025;36(3):36111-36128. https://doi.org/10.1088/1361-6501/adb06b

[16] Dou S, Dai F, Du Y. Deep Learning-Based Fault Diagnosis of Gearbox Using CBAM-DSC-ResNeXt50 Structure. Journal of Vibration Engineering & Technologies. 2025;13(6):418-422. https://doi.org/10.1007/s42417-025-01973-9

[17] Liang W, Zhang Z, Yin T, Zeng Y, Zhang Y. Multi-parallel disassembly line balancing problem and improved ant lion optimizer for mixed-waste electrical and electronic equipment. International Journal of Precision Engineering and Manufacturing-Green Technology. 2024;11(1):243-258. https://doi.org/10.1007/s40684-023-00525-4