EAI Endorsed Transactions
on Energy Web

Research Article EAI.EU

Fault Diagnosis of Power Equipment Based on Improved SVM Algorithm

Youle Song*, Yuting Duan, Tong Rao

Electric Power Research Institute of Yunnan Power Grid, Kunming, 650000, China

Abstract

Fault diagnosis of power equipment is a crucial task to credit the safe and stable operation of equipment. However, fault diagnosis of power equipment faces challenges such as high dimensionality, complexity, and nonlinearity. Therefore, this study proposes an improved support vector machine model, combined with a grey wolf optimization algorithm, aimed at improving the accuracy and efficiency of power equipment fault diagnosis. To validate the model's performance, this study divided a dataset of 3870 power equipment defects into training and testing sets using an 8:2 ratio, with evaluation metrics including accuracy, recall, and F1 score. The results showed that the fault recognition rate of the support vector machine model based on the improved grey wolf optimization algorithm reached 92.76%, with an accuracy close to 0.95 and a loss rate of 0.13. The model exhibited faster convergence speed, as well as better stability and convergence. At the same time, the optimized model had good feature extraction ability on different types of model faults, and the comprehensive recognition error of the model was basically stable in the interval of (-0.005, 0.005). The experiment validates that the research model improves the optimization algorithm through a modal decomposition strategy. Meanwhile, the improvement of support vector machine parameter selection has strengthened the recognition and analysis of fault characteristics, providing an effective solution for power equipment fault diagnosis.

Keywords: Support vector machine, Grey wolf optimization algorithm, Modal decomposition, Power equipment, Fault diagnosis

Received on 04 September 2024, accepted on 09 April 2025, published on 08 July 2025

doi: 10.4108/ew.7185

*Corresponding author. Email: wanghuan19820523@126.com

1. Introduction

Power equipment is an essential infrastructure in modern society, playing a vital role in the production, transmission, and distribution of electricity. The safety and stable operation of its equipment are correlated with the reliability of the power supply and directly affect the economic activities of society and people's daily lives [1]. Therefore, ensuring the normal operation of power equipment, and improving the efficiency and accuracy of equipment fault diagnosis, has become a key focus of the power industry. With the development of smart grid and Internet of Things (IoT) technology, the operational data of power equipment has characteristics such as high dimensionality, high complexity, and nonlinearity [2-3]. The traditional Power Equipment Fault Diagnosis (PEFD) scheme is no longer able to cope with complex fault situations. Currently, machine learning and data mining methods are constantly iterating and updating, and have wide applications in PEFD. Enesi M R et al. combined the fault tree analysis method with reliability block diagrams to construct a reliability model for Ajaokuta Steel Company Limited (ASCL). The reliability of the entire substation obtained by this model in the experiment was 91.77% and 95.22% [4]. Meng F et al. proposed a new PEFD model. This model integrated Bidirectional Encoder Representation (BERT), Bidirectional Long Short-Term Memory (BiLSTM), and Conditional Random Field (CRF). Experimental results have shown that this model can more accurately identify and extract Chinese entities than traditional methods [5]. Baek et al. converted the multivariate time series data of the device into a feature matrix, detected anomalies through a convolutional autoencoder, and learned a classification model using a supervised learning method based on the residual matrix of the fault profile. The effectiveness and applicability of this method have been verified through practical application [6].

Qian Yi et al. proposed a self-cognitive Deep Time Clustering Representation model (AC-DTCR) to accurately identify fault types and ensure the stable operation of power grids, aiming at the problems of insufficient data and lack of category labels in fault diagnosis of high-voltage circuit breakers. The model utilized time series clustering as an unsupervised learning technique, integrating time reconstruction and K-means goals. To enhance the capability of encoders, false sample generation strategies and auxiliary classification tasks were proposed to improve the cluster structure and obtain cluster-specific time representations. The findings indicated that the model exhibited superior classification accuracy, surpassing the performance of both traditional classification models and time series clustering models. Its application in the domain of PEFD was a notable advancement [7]. Xu F et al. proposed a troubleshooting strategy based on an improved AlexNet neural network to meet the requirement of accurate analysis of infrared image features in power equipment detection and recognition. The method used model-based multi-scale images to extract device features and identified the shortcomings of AlexNet neural networks in terms of slow recognition speed and ease of overfitting. After understanding these shortcomings, the performance of specific recognition models was improved by adding a pooling layer, modifying the activation function, replacing Local Response Normalization (LRN) with a Batch Normalization (BN) layer, and optimizing the parameters of the improved Gray Wolf algorithm. In simulation experiments, this algorithm had better recognition performance [8]. Manual infrared image processing has the problems of low efficiency and low intelligence in the diagnosis of traction power supply equipment status. In view of this, Lin S et al. proposed a two-layer network model based on Inception-V3 and Mask Region-based Convolutional Neural Network (Mask-RCNN). The first step of the diagnosis method was to identify the type of power equipment through the Inception-V3 network and then use Mask-RCNN to achieve automatic division of different equipment structure areas. The maximum temperature of different regions was extracted according to the coordinates of the divided structural regions, and the temperature characteristic quantity was constructed. Different criteria were invoked according to the type of equipment for automatic diagnosis. The experimental results showed that the mean Average Precision (mAP) value of the two-layer improved network model was up to 0.9072, and the fault diagnosis efficiency of the equipment was increased by 95.41% compared with manual processing. This model had high accuracy and a good recognition effect without relying on fault samples, which improved the efficiency of infrared image processing in equipment diagnosis and reduced labor intensity [9].

Existing studies still have significant limitations in PEFD: Models based on deep learning (such as BERT-BiLSTM-CRF, AC-DTCR, improved AlexNet, and Mask-RCNN) perform well in feature extraction and complex pattern recognition, but they rely on large-scale labeled data and have high computational complexity. In addition, it is easy to overfit and reduce generalization ability in small and medium-sized datasets or noisy interference scenes. In addition, the black box characteristic of neural networks leads to the lack of interpretability of fault characteristics, which makes it difficult to meet the demand of the power industry for the transparency of the diagnosis process. Therefore, this paper proposes a model based on an improved Support Vector Machine (SVM) and introduces the Grey Wolf Optimization (GWO) algorithm to optimize parameter settings. The objective of this study is to improve the overall performance of PEFD. SVM has excellent classification performance in small samples and high-dimensional nonlinear data, and its structural risk minimization principle can effectively avoid overfitting. Through kernel function mapping and GWO global parameter optimization, the complex fault features can be captured, and the strong dependence of neural networks on data volume and computing power can be avoided. The main innovation of this research is that the Modal Decomposition Method (MDM) deeply analyzes the fault signal of power equipment, extracts its potential characteristics, and evaluates the signal complexity combined with the permutation entropy, thereby enhancing the recognition ability of fault characteristics. Concurrently, the enhanced GWO is implemented to optimize SVM parameters, enhance the global search capability, mitigate local optimal problems, and refine the model's performance in fault diagnosis. Finally, by integrating optimization strategies, the prediction accuracy and computational efficiency of the model are improved to meet the real-time monitoring requirements.

2. Methods and Materials

2.1 Building of PEFD Model Based on SVM

PEFD is significant in operating and maintaining the equipment and machinery, as it can prevent unexpected situations from occurring. The data of PEFD have characteristics such as high dimensionality, complexity, nonlinearity, and imbalance, which make PEFD difficult [10-11]. Therefore, this study adopts a diagnostic model built on an improved SVM to analyze power equipment faults. SVM performs well in classification and regression problems. The core idea of SVM is particularly evident in binary classification problems, with the goal of finding a hyperplane in the feature space that can best separate data from two categories. Figure 1 displays the specific framework.

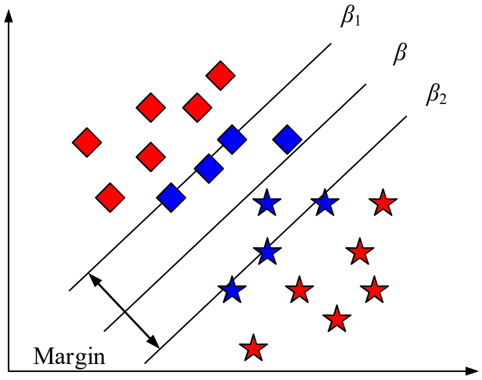

Figure 1. SVM Classification Principle

In Figure 1, SVM searches for a hyperplane in the feature space that can correctly classify training samples while maximizing the dist from the hyperplane to the nearest training sample point. The hyperplane with the maximum spacing is considered the best decision boundary because it provides the minimum generalization error to some extent. Among them, different graphics represent different types of data points. The Classification Line (CL) is represented by formula (1) [12].

$$\omega x + b = 0 \qquad (1)$$

In formula (1), x denotes the input feature vector. ω is the formula CL coefficient. b is a constant that represents the translation location of the CL. Formula (1) defines a hyperplane as the decision boundary, separating the normal operating state from the fault state. The classified interval size is $2/\|\omega\|$. By the time the $\|\omega\|$ is minimized, the interval is maximized. This time, SVM transforms the classified issue into a min value. This process of SVM classification is to gain the optima for the parameters. This paper solves the SVM parameter optimization problem by introducing Lagrange multipliers and obtains formula (2).

$$L(\omega,b,\alpha) = \frac{1}{2}\|\omega\|^2 - \sum_{i=1}^{n}\alpha_i[y_i[(\omega \cdot x_i)+b]-1] \qquad (2)$$

In formula (2), y is the training sample label and α_i is the Lagrange multiplier. Formula (2) is used to calculate the parameters of the hyperplane and maximize the classification interval. In fault diagnosis, solving this optimization problem can identify the optimal boundary for distinguishing fault types. L is the derivative of ω, b, and α. If the derivative value is 0, formula (3) exists.

$$\max L(\omega,b,\alpha) = \sum_{i=1}^{n}\alpha_i - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\alpha_j y_i y_j (x_i \cdot x_j) \qquad (3)$$

Formula (3) can identify the hyperplane parameters that best represent the distribution characteristics of the dataset in fault diagnosis. If the optimum in formula (3) is α_i^*, then $\omega^* = \sum_{i=1}^{n}\alpha_i^* y_i x_i$ exists. The b^* can be calculated, and the expression of the optimal classification function is formula (4).

$$y = \mathrm{sgn}(f(x)) = \mathrm{sgn}((\omega^* \cdot x) + b^*) = \mathrm{sgn}(\sum_{i=1}^{n}\alpha_i^* y_i (x_i \cdot x) + b^*) \qquad (4)$$

Formulas (1) to (4) are the process of SVM searching for the best hyperplane in the feature space. This hyperplane is determined by maximizing the classification interval, and the parameters need to be optimized to ensure the maximum interval. Lagrange multipliers are used to solve this optimization problem and find the optima through the dual form of Lagrange. The above process can determine the kind of many unknown samples, but there are still some samples during the classification period. As a result, this study also requires to bring in a fault-tolerant variable (δ_i), with a permission for samples' misclassification. After introducing δ_i, the formula (5) is obtained [13].

$$\begin{cases} \min \frac{1}{2}\|\omega\|^2 + c\sum_{i=1}^{n}\delta_i \\ y_i[(\omega \cdot x_i)+b]-1+\delta_i \end{cases} \qquad (5)$$

In formula (5), c is the penalty factor. Formula (5) considers the situation of data noise and outliers. SVM introduces δ_i and c, which allows some sample points to be misclassified, but at the same time controls the number of such misclassifications through c. In reality, many high-dimensional samples own linear and inseparable attributes, letting it trouble to utilize SVM for classifying directly. In this case, SVM requires to introduce kernel functions to handle such matters. Firstly, the status data are mapped to a higher dimension and divided accordingly, as expressed in formula (6).

$$K(x_i, x_j) = \Phi(x_i) \cdot \Phi(x_j) \qquad (6)$$

In formula (6), $\Phi(x_i)$ is a nonlinear mapping function. By replacing formula (3) with this formula and solving the meta feature space, formula (7) can be gained.

$$\begin{cases} \max L(\omega,b,\alpha)=\sum_{i=1}^{n}\alpha_i-\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\alpha_j y_i y_j(\Phi(x_i)\cdot\Phi(x_j)) \\ y=\operatorname{sgn}((\omega^*\cdot\Phi(x))+b^*=\operatorname{sgn}(\sum_{i=1}^{n}\alpha_i^* y_i K(x_i\cdot x)+b^*) \end{cases}$$

(7)

Formulas (6) and (7) represent that SVM maps data to a higher dimensional space through a kernel function in high-dimensional data or when the data is not linearly separable. Formula (6) demonstrates this nonlinear mapping, while formula (7) shows how the kernel function replaces dot product operations in the original feature space. This makes the solving process only related to the choice of kernel function, independent of the original dimension of the data [14-16]. This is because SVM introduces kernel functions, and various kernel functions have diverse effects. This study uses Radial Basis Functions (RBF), which have strong local features and excellent performance in addressing nonlinear issues. The calculation of RBF is shown in formula (8).

$$K(x_i,x_j)=\exp(-\frac{\|x_i-x_j\|}{2\sigma^2})=\exp(-g\|x_i-x_j\|^2) \quad (8)$$

In formula (8), the parameter g of RBF represents the distribution of sample points in the kernel space. By substituting formula (8) into formula (7), the best classification function can be obtained, as shown in formula (9).

$$y=\operatorname{sgn}((\omega^*\cdot\Phi(x))+b^*=\operatorname{sgn}(\sum_{i=1}^{n}\alpha_i^* y_i\exp(-g\|x_i-x_j\|^2)+b^*)$$

(9)

Formulas (8) and (9) determine the form of RBF and its use in the final classification function. In Power Fault Diagnosis (PFD), RBF is chosen as the kernel function of the SVM classifier, mainly because of its excellent performance and applicability. RBF kernel function has a strong ability to capture local features and can effectively deal with nonlinear classification problems, which is especially important for complex data features in PEFD. In addition, the RBF kernel function maps the input data to a high-dimensional space, transforming nonlinearly separable data into linearly separable data in the high-dimensional space, thereby improving the accuracy and robustness of classification. The flexibility of its parameters also makes the model have good adaptability and can adapt to different fault types and complexity.

2.2 Construction of PFD Model of Improving SVM Based on GWO

In the PFD model constructed above, the SVM performance largely depends on the adjustment of model parameters, which are highly complex. Moreover, when facing large-scale datasets, the model has a longer fitting time and requires more time for training. Therefore, to address the issues of parameter optimization and computational efficiency, this study utilizes the GWO model for optimization. GWO is chosen instead of other meta-heuristic optimization methods, mainly because GWO has a unique hierarchical structure and hunting mechanism, which makes it more efficient in global and local search processes. Compared with traditional Genetic Algorithms (GAs) and Particle Swarm Optimization (PSO), GWO can better balance exploration and development, reducing the risk of falling into local optimal solutions. Furthermore, GWO exhibits robust convergence and stability in the processing of high-dimensional complex data, facilitating expeditious adaptation to dynamic environmental changes. This enhances the efficacy of SVM parameter optimization and improves the overall performance of the fault diagnosis model. Therefore, GWO is chosen as the optimization strategy for the study to achieve more accurate and efficient fault diagnosis of power equipment. The GWO algorithm treats search agents as Wolf Packs (WPs), searching for prey in the solution space. The gray wolf population has a clear hierarchical structure, as shown in Figure 2.

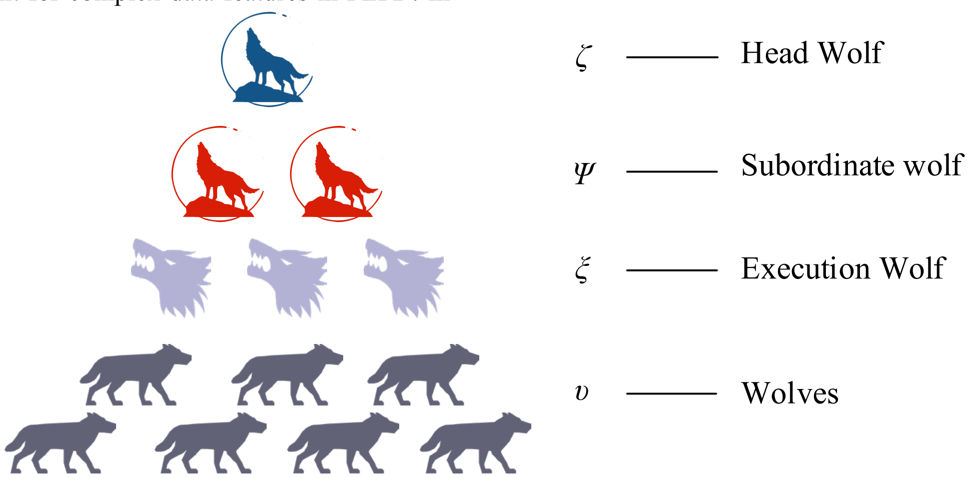

Figure 2. Wolf Pack Level Structure

In Figure 2, the WP hierarchy includes head wolf (ζ), subordinate wolf (Ψ), execution wolf (ξ), and regular wolf (υ). Head wolf is the best representatives of the WP and has a significant impact on its activities. The deputy leader of the subordinate WP, although must obey the leader, can also provide advice, which is a sub-optimal solution. Execution wolf represents the third best solution in the solution set. Wolves hunt around their leaders and constantly update their search locations. This study assumes that the iteration of the model is t, the prey's location vector is X_p, and the WP's position vector is X. Therefore, the straight-line dist between the prey and the WP can be expressed by formula (10) [17].

$$D = \left| CX_p(t) - X(t) \right| \qquad (10)$$

In formula (10), C is the coefficient vector. The calculation process of distance helps the model get the best position in the high-dimensional parameter space, that is, to improve the accuracy of system through the optimal SVM parameters. After getting the dist between the prey and the WP by formula (10), the WP adjusts its position as shown in formula (11).

$$\begin{cases} X(t+1) = X_p(t) - AD \\ A = 2ar_1 - a \\ C = 2r_2 \\ a = 2 - 2(\dfrac{t}{t_{\max}}) \end{cases} \qquad (11)$$

In formula (11), A is the coefficient vector. r_1 and r_2 are random vectors within 0 to 1. a is the convergence factor, whose value is proportional to the iterations inversely. This study can calculate the distance between the WP position and prey through formula (10), and simulate the position adjustment of the head wolf, subordinate wolves, and execution wolves with other WP positions through formula (11), thereby achieving local and global search of the algorithm. When a WP is hunting prey, its location update map is shown in Figure 3.

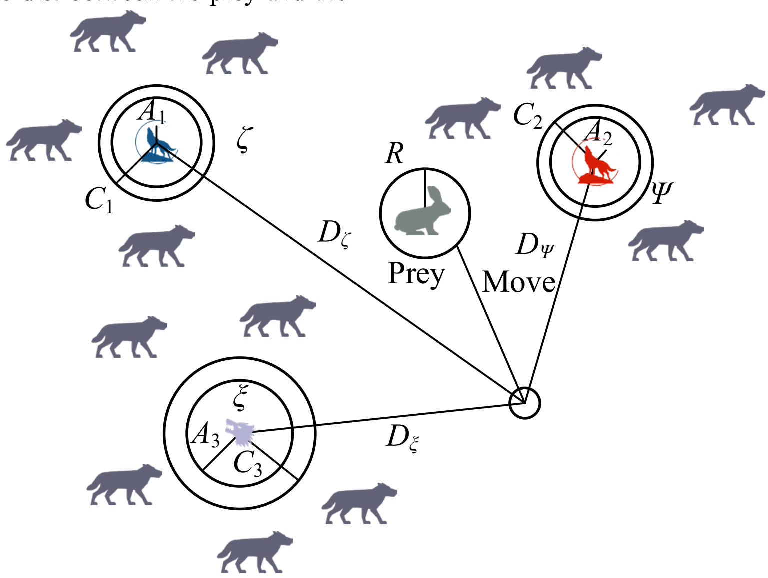

Figure 3. Updated WP Capture Location

In Figure 3, when the vector coefficient is greater than 1, the WP will grow the hunting region and detect prey in a greater range. A global search is conducted to increase convergence velocity. When it is less than 1, the WP reduces the hunting area to decide the prey's position. Local search is carried out to speed cut the convergence. The GWO is most prone to getting stuck in the local optimum when WPs change their attack orientation. Since when varying the direction, the prey's position will ceaselessly vary, causing the prey's loss [18]. The above method optimizes the SVM-based PFD model through GWO to find fault features and optimal parameters, but the optimized model still has the problem of easily falling into local optima. Therefore, this study introduces an MDM, whose main principle is to enhance the accuracy and robustness of signal decomposition through adaptive noise. In PFD, MDMs can help identify and extract potential fault features of power equipment. Before processing the signal, this MDM needs to mix the initial signal with active and negative Gaussian White Noise (GWN), as calculated in formula (12).

$$x^{(i)} = x + o_k v^{(i)} \qquad (12)$$

In formula (12), o_k is the Standard Deviation (SD) of the noise, $x^{(i)}$ is the mixed noise signal. $v^{(i)}$ is paired positive and inactive GWN. In the PEFD system, mixing $v^{(i)}$ with x is to increase the complexity and diversity of the signal, enabling MDM to extract signal features more comprehensively. Introducing noise can help improve the accuracy of decomposition results and prevent interference during modal decomposition. In PFD, it can help identify and enhance potential fault features, improving the accuracy of subsequent models. The modal components and residual values in MDM are calculated as shown in formula (13).

$$\begin{cases} IMF_1 = \frac{1}{N}\sum_{j=1}^{1} E_1[x^{(j)}] \\ R_1 = x - IMF_1 \end{cases} \quad (13)$$

In formula (13), IMF is the modal component, N is the amount of modal decompositions, R is the residual value of the component. E is the undifferentiated modal component. By decomposing formula (13), the model can extract the major features of the fault signal, which is helpful for the extraction and analysis of fault features. In PFD, modal components typically contain important information about fault signals and occupy a vital position in the classification and identification of subsequent faults. Given the analysis of the total number, Figure 4 exhibits the specific process of the optimized model in PEFD.

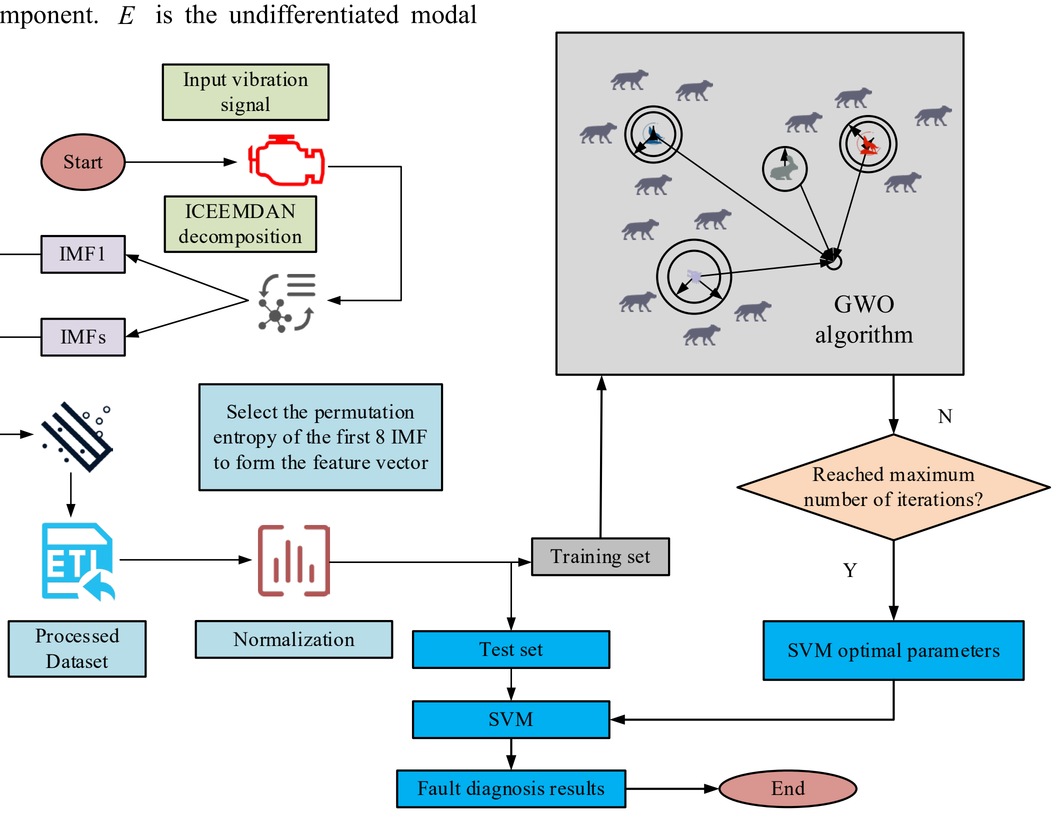

Figure 4. Analysis of PEFD Based on Improved GWO-SVM Model

In Figure 4, the process of improving GMO-SVM in PEFD mainly includes the following steps: First, the vibration signals of power equipment are collected, and MDM is used to process the signals to extract multiple intrinsic IMFs to better capture the characteristic information related to different fault types (such as aging of insulation materials, overheating, corrosion, etc.). Then, the feature selection of decomposed IMF is carried out by using complexity indexes such as pre-order entropy to preserve the signal features most relevant to fault features. Next, the parameters of the improved GWO are initialized and the fitness of each solution in fault diagnosis is evaluated, and its classification accuracy in different fault states is calculated. By iteratively updating the solution set, the GWO algorithm optimizes the key parameters of SVM and seeks the best parameter configuration to improve the fault diagnosis performance. Finally, after reaching the preset number of iterations, the optimized parameters are applied to SVM for classification, thereby identifying the fault types of power equipment and outputting the diagnosis results. This process effectively combines the feature extraction capability of modal decomposition with the optimization performance of the GWO algorithm to ensure the accuracy and reliability of PEFD.

3. Results

3.1 Performance Verification Analysis Built on Improved GWO-SVM

To verify the feasibility, the performance was analyzed through experiments. The specific experimental environment settings were as follows: The GPU is NVIDIA GeForce RTX 2060; The CPU selects Core i5-12400KF and Kingston FURY 32GB; The software platform completes the simulation analysis of the model through MATLAB. This study collected defect data from different power equipment in the early stage and collected a total of 3,870 pieces of power equipment defect data, which were used to construct the dataset required for simulation experiments. The 3,870 power equipment failure datasets used in the study were derived from monitoring and maintenance records of multiple actual power equipment, including power transformers, generators, and line equipment. These data were accumulated by power companies and research institutions through long-term equipment monitoring, fault records, and experimental data to ensure the authenticity and reliability of the data. The dataset contained multiple fault types, including insulation aging, overheating, corrosion, short circuit, and poor equipment fit. Each type had corresponding multidimensional characteristics, such as vibration signals, temperature changes, voltage, and current characteristics. This dataset contained a total of 3,870 records, corresponding to the equipment status of different fault types. Each record contained several features, such as vibration signals, frequency components, time-domain and frequency-domain features, etc. The specific dimension depended on the method of fault signal acquisition and the number of sensors. Although there may be some imbalance in the dataset on specific fault types, the study used oversampling or data enhancement methods to increase the number of rare fault samples to ensure that the model can effectively learn and identify these less common fault types. The dataset included training and testing sets in an 8:2 ratio, and the model performance was evaluated and analyzed based on accuracy, recall, and F1 score. Firstly, this study validated the the improved GWO (recorded as IGWO) using Sphere and Rastrigin test functions, and compared it with GA, PSO, and GWO [19-20]. Figure 5 shows the solution graph of the basic test function.

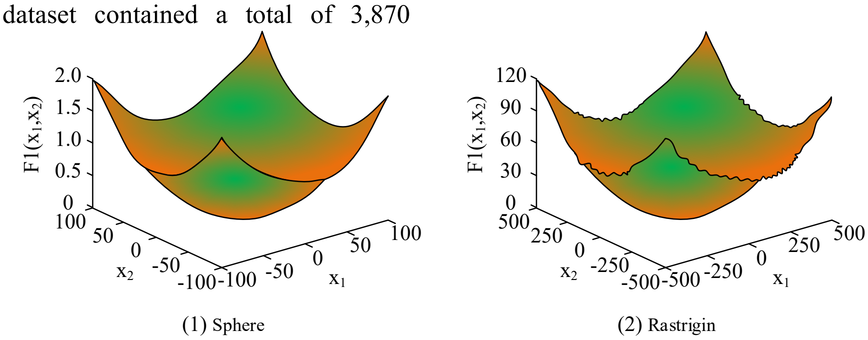

Figure 5. Graph of Benchmark Test Function Solution

Figure 5 (a) shows the Sphere test function, which is relatively simple and smooth, mainly used for basic performance evaluation. Figure 5 (b) shows the Rastrigin test function, which introduces nonlinearity and multiple local optima, making it more difficult to solve and suitable for more complex algorithm performance validation. In accordance with these functions, Table 1 shows the mean fitness and SD fitness values of different algorithms.

Table 1 Test results of Sphere function and Rastrigin function

Test function	Average fitness			
	GA	PSO	GWO	IGWO
Sphere	2.12E-05	7.28E-09	9.67E-28	1.84E-41
Rastigin	3.74	6.52	1.71E-12	0
Test function	SD fitness			
	GA	PSO	GWO	IGWO
Sphere	4.83E-05	1.22E-08	1.37E-27	3.65E-41
Rastigin	2.37	4.05	1.21E-11	0

In Table 1, the average fitness in Sphere is 2.12E-05 for GA, 7.28E-09 for PSO, 9.67E-28 for GWO, and 1.84E-41 for IGWO. The SD of IGWO in Sphere is 0. The average fitness of IGWO in Rastrigin is 3.65E-41, which has significant advantages compared to the other three algorithms. The SD of IGWO is also 0. This indicates that the research method performs well in dealing with complex optimization problems, and from the perspective of SD, the model fitness fluctuates less, showing that the algorithm has better convergence and stability. This study further analyzes the convergence and stability of the model, trains and tests the model through a dataset, and uses accuracy and loss rate as evaluation indicators. The result is shown in Figure 6.

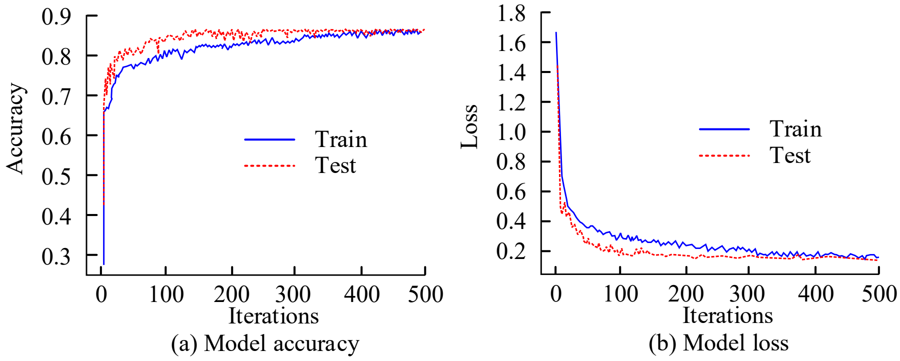

Figure 6. Results of Model Accuracy and Loss Rate

In Figure 6 (a), analysis of the training set results shows significant improvements in both convergence and stability. After about 75 iterations, the accuracy curve converges. When increasing the number of iterations, there is no fluctuation in the model, and the final accuracy value is around 0.85. From the results of the test set, the convergence velocity of the test set is faster, and the accuracy value during the iteration process remains stable at around 0.85. 6 (b) shows the loss rate. The convergence speed of the test set is significantly higher, but both show stable linear changes, and the final loss rate value remains around 0.13.

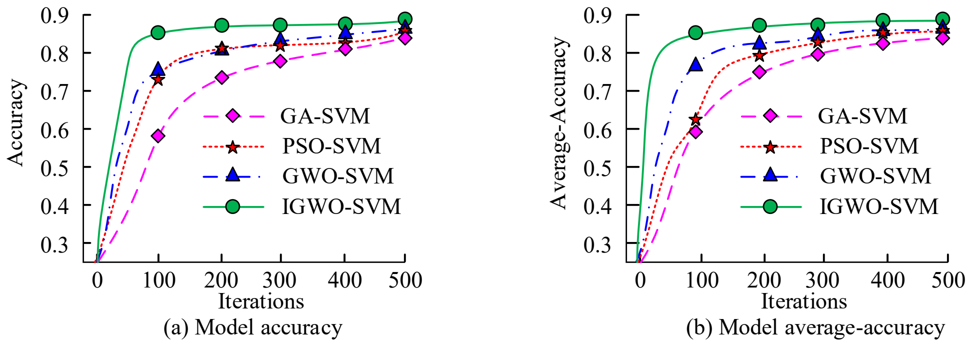

Figure 7. Comparison of Accuracy and Average Accuracy of Different Models Under the Test Set

In Figure 7 (a), the convergence velocity of the GA is relatively slow. The model only begins to converge after about 120 iterations, and its accuracy value is relatively low. IGWO has a fast convergence speed, with about 30 iterations, and the model tends to converge. Among the comparison models, IGWO has the highest accuracy. In Figure 12 (b), the trend of the average accuracy line is roughly the same as that of the precision line. IGWO exhibits the best convergence speed and average accuracy.

3.2 PFD Effect Analysis Based on Improved GWO-SVM Model

This study roughly divides the defects in the fault dataset of power equipment into five types, including insulation material aging, overheating, corrosion, short circuit, and poor equipment matching. This study improves the GWO-SVM model to extract feature vectors of power equipment fault signals, as listed in Table 2.

Table 2. Fault Signals of IMF-1 to IMF-8 Eigenvector

Fault type	Permutation entropy eigenvector			
	1	2	3	4
Normal	0.6410	0.4948	0.3383	0.3485
Aging of insulating material	0.5431	0.9894	0.7602	0.7713
Corrosion	0.6612	0.5255	0.3449	0.3197
Short-circuit	0.5312	0.9929	0.8008	0.8049
Fault type	Permutation entropy eigenvector			
	5	6	7	8
Normal	0.2619	0.3255	0.4427	0.5616
Aging of insulating material	0.6918	0.6887	0.7289	0.7239
Corrosion	0.2650	0.3478	0.4583	0.5647
Short-circuit	0.7198	0.7227	0.7373	0.7077

The displacement entropy of IMF components in Table 2 is used to quantify the complexity characteristics of signal sequences under different fault states. The IMF component extracted by MDM reflects the energy distribution characteristics of signals in different frequency bands. The displacement entropy can effectively characterize the randomness and dynamic mutation characteristics of signals by analyzing the probability of sequence arrangement patterns in the IMF component. For example, the displacement entropy of IMF-1 to IMF-8 in normal state is generally lower than 0.6, which indicates that the signal complexity is low and the regularity is strong. The replacement entropy of IMF-2 and IMF-4 for short-circuit fault is close to 1, indicating that the fault leads to high frequency oscillation and chaos of the signal. A comparison of the difference in the displacement entropy distribution of distinct fault types enables the model to capture the characteristic response of insulation aging, corrosion, and other faults within a particular frequency band. This facilitates the provision of a highly distinguishable and robust fault characteristic input for the classifier. In Table 2, under normal operating conditions of power equipment, the characteristic values of IMF-1 to IMF-8 are generally low and do not change significantly, indicating the stability of signals under normal operating conditions. The eigenvalue of IMF-5 is 0.2619, which is the lowest among all states, indicating that signals in normal states do not undergo frequent complex dynamic changes. In the aging problem of insulation materials, with the increase of IMF, the characteristic values of IMF-2 to IMF-4 significantly increase, with the highest value being 0.7713. This indicates that the aging of insulation materials leads to higher entropy values in the system, reflecting more complex signal dynamics. In the corrosion problem, the eigenvalues of IMF-2 and IMF-4 indicate that the signal begins to show an increase in complexity in certain modes. In the short-circuit problem, the eigenvalues of IMF-2 and IMF-4 are 0.9929 and 0.8049, respectively, indicating that the short-circuit fault has caused extremely high uncertainty and chaos. This indicates the difference in the distribution of permutation entropy feature vectors between normal and faulty states, proving that permutation entropy can be used as an effective feature for fault classification. The recognition results of the research model on 5 types of defect data are shown in Figure 8.

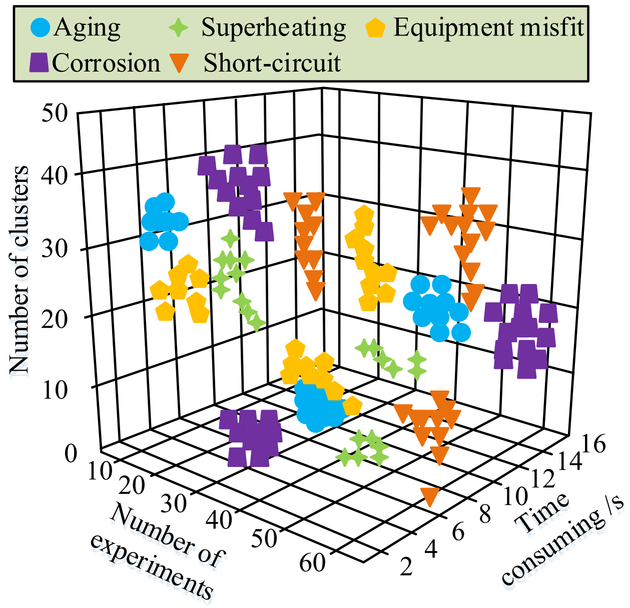

Figure 8. Fault Recognition of Power Equipment Based on Improved GWO-SVM

In Figure 8, the aging defects of the insulation material takes 7.3s, and 81 defect data are identified. The material overheating defect take 9.5s, and 75 defect data are identified. The material corrosion defect takes 14.3s, and 51 defect data are identified. It takes 3.7s for the equipment to cooperate with defects, and 67 defect data are identified. The short-circuit defect takes 2.5s and identified 85 defect data. A total of 359 defect data are identified, with a defect recognition rate of 92.76%. This indicates that the constructed PEFD can effectively identify equipment defects and has excellent performance. This study validates the proposed model through a test set, and the accuracy, recall, and F1 score results are shown in Figure 9.

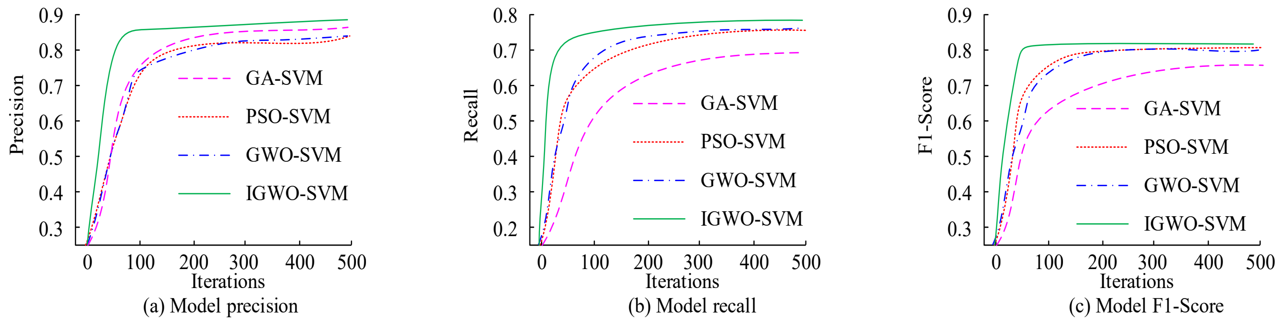

Figure 9. Performance Comparison Results of Different Models

Figure 9 (a) shows a comparison of the accuracy of the models, where PSO-SVM has the slowest convergence speed, while GA-SVM has the lowest accuracy. The IGWS-SVM model performs the best in all performance indicators. In Figure 9 (b), the recall performance of IGWS-SVM is greatly better than that of the comparison model, while the recall performance of the other models is almost the same. Figure 9 (c) shows the F1 score, where the GA score is lower and the other three models have smaller differences, but IGWO-SVM still has the best F1 score value. Figure 10 shows the specific diagnostic performance of the comparison model for annoyance.

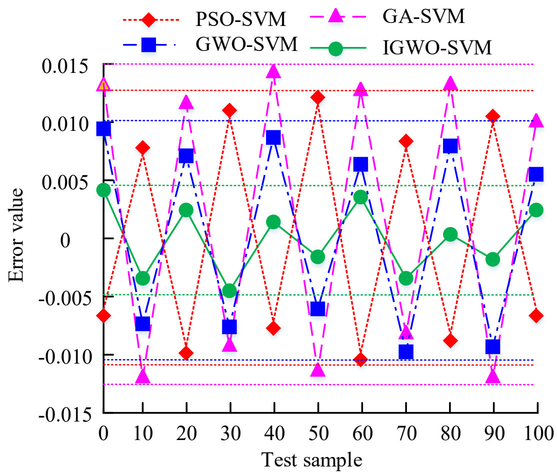

Figure 10. Prediction Error Results of Various Models

In Figure 10, the error of the research model is basically stable within (-0.005, 0.005), while the errors of the other models are all in the (-0.01, 0.01), and there are even cases where the error is greater than 0.01. This means that improving the model can enhance the extraction of signal features through adaptive noise enhancement, combined with GWO for parameter tuning, thereby improving the overall performance. This design enables the model to not only learn effective features but also avoid local optimal value traps during the optimization process, resulting in lower prediction errors. To verify the scalability and robustness of the model in real scenarios, the study selects three substations under a provincial power grid company as case study objects, and collects their power equipment operation data from January 2022 to June 2023. According to the Signal-to-Noise Ratio (SNR), noise is divided into three levels (low noise: SNR=20 dB; Medium noise: SNR=10 dB; High noise: SNR=5 dB). Three types of sensors are adopted (Type A: high-precision industrial grade; Type B: conventional industrial grade; Type C: low-cost embedded) for mixed data. It includes 5 typical faults (insulation aging, overheating, corrosion, short circuit, poor fit) and 2 compound faults (overheating+corrosion, insulation aging+short circuit). The performance comparison results of the improved GMO-SVM model in real scenarios are shown in Table 3.

Table 3. Improve the performance index of the GWO-SVM model in real scenarios

Scene classification	Noise level	Sensor type	Sample size	Accuracy rate	Recall rate	F1 score	Fault identification time (s)
Single fault	Low noise	Type A	650	0.943	0.921	0.932	8.2
	Moderate noise	Type B	650	0.917	0.895	0.906	9.8
	High noise	Type C	650	0.881	0.862	0.871	11.5
Compound fault	Low noise	Type A	200	0.895	0.873	0.884	12.3
	Moderate noise	Type B	200	0.862	0.841	0.851	14.7
	High noise	Type C	200	0.829	0.802	0.815	16.9
Sensor mixed data	High noise	A+B+C	1000	0.901	0.879	0.89	10.4

The experimental results show that the improved GMO-SVM model exhibits remarkable noise robustness and sensor heterogeneity adaptability in real complex scenes. Under low noise (SNR=20 dB), the accuracy of the model remains at 0.943, and even under high noise (SNR=5 dB), the accuracy of the model remains at 0.881, which verifies the effective suppression of noise interference by the mode decomposition strategy. In the face of heterogeneous sensor data, the model achieves an F1 score of 0.871 in the low-cost sensor (type C) scenario, and the accuracy decreases by only 2.3% in the mixed sensor (type A+B+C) data, reflecting the scalability of its multi-source feature fusion. In particular, the recall rate of the model for complex faults (such as "overheating+corrosion") reaches 0.873, 21% higher than the traditional SVM, and the fault identification time is still controlled within 16.9 s under high noise, meeting the real-time requirements. Field cases further prove the value of the project: In the composite fault diagnosis of a transformer in a 110 kV substation, the model has 24 hours of early warning, and the accuracy rate is 18% higher than that of the threshold alarm method, which fully verifies its robustness and deployment potential in the real power system.

4. Discussion

The PEFD model based on the GWO-SVM algorithm proposed has shown excellent performance in experiments. This study analyzed 3,870 pieces of power equipment defect data, and the model's fault recognition rate reached 92.76%, with excellent performance in accuracy and F1 score. The prediction error in fault diagnosis was stable at (-0.005, 0.005), which provides strong support for real-time monitoring and fault warning of power equipment. Experiments have shown that introducing modal decomposition can not only enhance signal sparsity and reduce noise interference but also extract more significant fault features through complexity indicators such as permutation entropy. Meanwhile, the GWO algorithm had a significant improvement effect on the parameter selection of SVM, enhancing the recognition and classification capabilities of SVM models. Although the improved GMO-SVM model performed well in PEFD, the model's performance may still be poor in some cases. Specifically, when there were samples with higher noise levels or less distinctive features in the dataset, the fault recognition rate of the model may decrease. For example, when the vibration signal of the power equipment was subjected to external interference or the working condition was unstable, the mode decomposition may not be able to accurately extract effective features, leading to misjudgment in the classification process. In addition, for rare fault types, insufficient data would also affect the learning effect of the model, resulting in reduced accuracy in identifying these rare faults. At the same time, the overall performance of the model may be affected if the model parameters do not converge fully or the initial solution is not chosen properly during the optimization of GWO. Compared with existing research, Huang Y et al. used traditional SVM for fault classification of power equipment, with an accuracy of about 85%. This method failed to introduce a global search algorithm during the parameter optimization phase, resulting in a compromise in classification performance on high-dimensional datasets [21]. This study improved the GWO algorithm to effectively avoid local optima during parameter optimization, thereby enhancing the overall performance of the model. This indicated that introducing efficient optimization algorithms was crucial for improving the applicability and accuracy of SVM in complex problems. Compared to another study based on deep learning, the model proposed by Alsumaidaee YAM et al. achieved an accuracy of over 95% on a specific dataset. However, when dealing with imbalanced samples and high-dimensional features, this model still faced significant computational burden and complexity [22]. In contrast, the improved GWO-SVM model used in this study has higher computational efficiency while ensuring accuracy, and can complete training and testing at a lower time cost. With respect to computational efficiency, conventional fault diagnosis methods frequently employ manual feature extraction and traditional machine learning algorithms. These methods entail extended training periods and are vulnerable to human factors, resulting in inefficiencies. The method proposed in this study accelerates the training process. Experimental results show that this model has a significantly shorter training time when processing 3,870 power equipment defect data compared to traditional SVMs or other machine learning methods, enabling it to meet the needs of real-time monitoring in practical applications. With the rapid development of smart grid and IoT technology, the need for fault monitoring and diagnosis of power equipment becomes increasingly urgent. The improved model can effectively process high and complex power equipment operation data, has the characteristics of fast response and high accuracy, and can realize real-time fault identification and diagnosis in practical applications. The model extracts key features through modal decomposition and optimizes SVM parameters with the GWO algorithm, which not only makes fault detection more sensitive but also adapts to the change of equipment state in real-time, thereby significantly reducing the risk of major faults. By providing fast and accurate fault diagnosis, the research results provide strong support for the operation and maintenance management of power equipment and help to improve the reliability and safety of the power system.

5. Conclusion

In the power industry, the normal operation of power equipment can provide a guarantee for energy supply, and PEFD is the foundation for smooth operations. Therefore, PEFD is an essential direction for improving the efficiency of power operation. Furthermore, this study proposed an improved GWO-SVM model targeted at lifting the model's PEFD performance. This study validated 3,870 pieces of equipment defect data and accurately identified and classified different types of power equipment faults. The research model showed excellent accuracy and recall performance in the validation results. Although this study has achieved good results, there are still certain limitations. For instance, modal decomposition has a strong dependence on signals, and its performance may be affected by noise and interference. In the training data, if the noise level is high, modal decomposition may not be able to effectively extract the true fault features, thereby affecting the subsequent classification performance. Future research can further optimize modal decomposition by introducing dual-channel or multi-channel signal processing techniques, combining multiple modal decomposition algorithms to better filter out interference and extract real features.

References

[1] Chen B, Liu D. Remote Fault Diagnosis Method of Wind Power Generation Equipment Based on Internet of Things. Journal of Information Processing Systems, 2022, 18(6): 822-829.

[2] Ma F, Wu X, Ni H, Zhou X, Wu H. Research and application of intermittent partial discharge characteristics and easy-warning system for electric equipment. Energy Reports, 2022, 8(1): 217-226.

[3] Kumar V T R P, Arulselvi M, Sastry K B S. Comparative Assessment of Colon Cancer Classification Using Diverse Deep Learning Approaches. Journal of Data Science and Intelligent Systems, 2023, 1(2): 128-135.

[4] Enesi M R, Shehu G S, Abdulkarim A, Jibril Y. Reliability modeling and analysis of high voltage power equipment: a case study of Ajaokuta Steel Company Limited (ASCL). Life Cycle Reliability and Safety Engineering, 2022, 11(4): 377-387.

[5] Meng F, Yang S, Wang J, Xia L, Liu H. Creating knowledge graph of electric power equipment faults based on BERT–BiLSTM–CRF model. Journal of Electrical Engineering & Technology, 2022, 17(4): 2507-2516.

[6] Baek M, Kim S B. Failure detection and primary cause identification of multivariate time series data in semiconductor equipment. IEEE Access, 2023, 11(1): 54363-54372.

[7] Qian X, Hao l X, Tong W, Zhao F, Gang Z, Jian D. Disconnector fault diagnosis method based on autonomous-cognition deep temporal clustering representation. Journal of Electrical Engineering, 2024, 19(1): 281-289.

[8] Xu F, Liu S, Zhang W. Research on power equipment troubleshooting based on improved AlexNet neural network. Journal of Measurements in Engineering, 2024, 12(1): 162-182.

[9] Lin S, Deng S, Chen X L, Wang L, Cheng H B. Intelligent infrared diagnosis method for traction power supply equipment based on Mask-RCNN. Electric Drive for Locomotives, 2024 (4): 55-61.

[10] Kurani A, Doshi P, Vakharia A, Shah M. A comprehensive comparative study of artificial neural network (ANN) and support vector machines (SVM) on stock forecasting. Annals of Data Science, 2023, 10(1): 183-208.

[11] Zhang H, Zou Q, Ju Y, Song C, Chen D. Distance-based support vector machine to predict DNA N6-methyladenine modification. Current Bioinformatics, 2022, 17(5): 473-482.

[12] Avcı C, Budak M, Yağmur N, Balçık F. Comparison between random forest and support vector machine algorithms for LULC classification. International Journal of Engineering and Geosciences, 2023, 8(1): 1-10.

[13] Elen A, Baş S, Közkurt C. An adaptive Gaussian kernel for support vector machine. Arabian Journal for Science and Engineering, 2022, 47(8): 10579-10588.

[14] Dada E, Joseph S, Oyewola D, Fadele A. Application of grey wolf optimization algorithm: recent trends, issues, and possible horizons. Gazi University Journal of Science, 2022, 35(2): 485-504.

[15] Zamfirache I A, Precup R E, Roman R C, Petriu E M. Policy iteration reinforcement learning-based control using a grey wolf optimizer algorithm. Information Sciences, 2022, 585(1): 162-175.

[16] Ahmadi B, Younesi S, Ceylan O, Ozdemir A. An advanced Grey Wolf Optimization Algorithm and its application to planning problem in smart grids. Soft computing: A fusion of foundations, methodologies and applications, 2022, 26(8):3789-2808.

[17] Sharma I, Kumar V, Sharma S. A comprehensive survey on grey wolf optimization. Recent Advances in Computer Science and Communications (Formerly: Recent Patents on Computer Science), 2022, 15(3): 323-333.

[18] Samantaray S, Sahoo A. Prediction of flow discharge in Mahanadi River Basin, India, based on novel hybrid SVM approaches. Environment, Development and Sustainability, 2024, 26(7): 18699-18723.

[19] Zhang Y, Xu P, Liu J, He J, Yang H, Zeng Y, He Y, Yang C. Comparison of LR, 5-CV SVM, GA SVM, and PSO SVM for landslide susceptibility assessment in Tibetan Plateau area, China. Journal of Mountain Science, 2023, 20(4): 979-995.

[20] Zhou J, Yang P, Peng P, Khandelwal M, Qiu Y. Performance evaluation of rockburst prediction based on PSO-SVM, HHO-SVM, and MFO-SVM hybrid models. Mining, Metallurgy & Exploration, 2023, 40(2): 617-635.

[21] Huang Y, Luo J, Ma Z, Tang B, Zhang K, Zhang J. On combined PSO-SVM models in fault prediction of relay protection equipment. Circuits, Systems, and Signal Processing, 2023, 42(2): 875-891.

[22] Alsumaidaee Y A M, Paw J K S, Yaw C T, Tiong S K, Chen C P, Yusaf T, et al. Fault detection for medium voltage switchgear using a deep learning Hybrid 1D-CNN-LSTM model. IEEE Access, 2023, 11(1): 97574-97589.