processes

MDPI

Article

A Data-Driven Strategy Assisted by Effective Parameter Optimization for Cable Fault Diagnosis in the Secondary Circuit of a Substation

Dongbin Yu [1], Yanjing Zhang [2], Sijin Luo [2], Wei Zou [2], Junting Liu [2], Zhiyong Ran [3] and Wei Liu [4,*]

[1] Guangdong Power Grid Limited Liability Company Jiangmen Power Supply Bureau, Jiangmen 529000, China; ydb-1@163.com
[2] Jiangmen Power Supply Bureau, China Southern Power Grid, Jiangmen 529000, China; 18128281872@163.com (Y.Z.); 13929055828@163.com (S.L.); nestchow@163.com (W.Z.); liujunting0213@163.com (J.L.)
[3] XJ Electric Co., Ltd., Xuchang 461000, China; 13849858531@163.com
[4] School of Electrical and Electronic Engineering, AnHui Science and Technology University, Bengbu 233100, China
* Correspondence: weiliuutsa@163.com

Abstract

As power systems evolve rapidly, cables, essential for electric power transmission, demand accurate and timely fault diagnosis to ensure grid safety and stability. However, current cable fault diagnosis technologies often struggle with incomplete feature extraction from complex fault signals and inefficient parameter tuning in diagnostic models, hindering efficient and precise fault detection in modern power systems. To address these, this paper proposes a data-driven strategy for cable fault diagnosis in substation secondary circuits, enhanced by effective parameter optimization. Initially, wavelet packet decomposition is employed to finely divide collected cable fault current signals into multiple levels and bands, effectively extracting fault feature vectors. To tackle the challenge of selecting penalty and kernel parameters in Support Vector Machine (SVM) models, an improved Golden Jackal Optimization (GJO) algorithm is introduced. This algorithm simulates the predatory behavior of golden jackals in nature, enabling efficient global optimization of SVM parameters and significantly improving the classification accuracy and generalization capability of the fault diagnosis model. Simulation verification using real cable fault cases confirms that the proposed method outperforms traditional techniques in fault recognition accuracy, diagnostic speed, and robustness, proving its effectiveness and feasibility. This study offers a novel and efficient solution for cable fault diagnosis.

Keywords: cable fault diagnosis; wavelet packet decomposition; support vector machine (SVM); improved golden jackal optimization (GJO) algorithm

check for updates

Academic Editor: Thair Mahmoud

Received: 25 May 2025
Revised: 8 July 2025
Accepted: 24 July 2025
Published: 29 July 2025

Citation: Yu, D.; Zhang, Y.; Luo, S.; Zou, W.; Liu, J.; Ran, Z.; Liu, W. A Data-Driven Strategy Assisted by Effective Parameter Optimization for Cable Fault Diagnosis in the Secondary Circuit of a Substation. *Processes* **2025**, *13*, 2407. https://doi.org/10.3390/pr13082407

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

With the rapid development of urban construction, the scale of power grids has become increasingly vast, and electric energy has become an indispensable source of energy in people's daily lives [1]. Cable lines play a vital role in electric power transmission and distribution, serving as the hub of power transportation [2]. Modern urban planning places greater emphasis on ecology and livability, requiring the planning and layout of transmission and distribution equipment to consider environmental and spatial needs. Due to their convenient installation, stable and reliable performance, and minimal space

occupation, power cables are increasingly occupying a larger proportion in the transmission network, gradually replacing overhead lines [3]. As cities continue to expand rapidly and electrical equipment is continuously put into use, the scale and specifications of cable laying continue to increase. As the core of urban power equipment, the reliable operation of power cables is crucial to urban development and residents' lives [4].

Power cables are generally laid underground, which makes it more difficult to locate and diagnose faults when they occur. Locating the fault points often consumes significant human, material, and financial resources. Additionally, cables lack the direct observability of overhead lines [5], making the task of fault detection and diagnosis for cables more arduous and challenging. Since fault diagnosis of power cables is crucial for the safe and stable operation of power systems, achieving rapid and accurate diagnosis of fault disturbances can promptly identify insulation defects and hidden dangers, prevent further deterioration of faults, and significantly reduce the loss of human, material, and financial resources caused by faults [6,7].

At present, the main methods for power cable fault diagnosis both domestically and internationally include DC superposition, DC component analysis, dielectric loss measurement, partial discharge detection, and infrared temperature measurement [8–10]. However, the accuracy of DC superposition and DC component analysis is not high, the applicability of dielectric loss measurement is limited, the signal collected by partial discharge detection is very weak, and the battery life of infrared temperature measurement instruments is short with high maintenance costs. These methods all need improvement in detection efficiency and accuracy. In recent years, many researchers have applied artificial intelligence algorithms and wavelet theory signal processing techniques to cable fault diagnosis, based on the characteristics of power cable transient signals and traditional fault diagnosis methods. Cable fault diagnosis methods at home and abroad mainly consist of two processes: transient signal feature extraction and classification processing [11–13]. Literature [14] realizes cable fault diagnosis based on input impedance spectroscopy, utilizing the power cable distributed parameter model, input impedance spectroscopy, spectral analysis, and the Kaiser window. By modeling different types of cable faults, it obtains the input impedance spectroscopy characteristics of cables under normal and faulty conditions to discriminate fault types. Neural networks and wavelet transforms have also been widely applied in fault diagnosis in recent years. Literature [15] first uses cable fault simulation data based on the MATLAB platform to train a deep neural network (DNN) model, which is then used to diagnose cable faults. In terms of intelligent algorithm training, experts and scholars have been exploring typical feature vectors that can constitute critical fault features as training template libraries, including wavelet energy spectrum, wavelet singular entropy, and average current zero-crossing rate [16,17], conducting numerous studies to improve the accuracy of cable fault diagnosis.

Literature [18] utilizes both approximate entropy and the standard deviation of integral values above and below the time axis to distinguish between non-fault lightning strikes, fault lightning strikes, and short-circuit faults. Based on the wavelet analysis results of transient current signals, Literature [19] directly utilizes a current transformer to extract transient high-frequency current features, calculate distribution weights, and achieve fault phase selection based on wavelet entropy weights of high-frequency transient components. Literature [20] proposes a fault recognition principle for different voltage phase angle intervals, using the time-frequency characteristics of wavelet transforms to extract the traveling wave energy of each phase-mode current after a fault for fault phase selection. Literature [21] constructs feature vectors using wavelet packet information entropy and inputs them into a trained CS-BP neural network to detect power cable fault types. Literature [22] fuses multi-source information using fuzzy theory, establishes a recognition framework

and evidence body based on D-S evidence theory, and determines the reasonable allocation of belief function values based on expert experience-assigned parameter weights. Finally, the maximum belief criterion is used to judge the operating state of the tested cable. Literature [23] performs multi-resolution analysis on current signals using wavelet transforms, extracts their energy entropy as input vectors for a probabilistic neural network (PNN) fault recognition model, and proposes a loose fault recognition method based on wavelet energy entropy and PNN. Literature [24] proposes a multi-scale analysis method for traveling wave signals based on wavelet transforms, which can effectively extract fault information and accurately locate fault positions. Literature [25] uses the one-dimensional Mallat algorithm of wavelet transforms to analyze traveling wave signals for determining cable fault locations. However, these studies only distinguish between short-circuit and lightning strike signals or perform fault phase selection, unable to diagnose multiple fault types and their causes in cables. Due to the complexity of the cable operating environment and significant interference from the fundamental frequency component of cable currents, ordinary time-frequency domain analysis is commonly used in current research, which cannot effectively extract current signal features, leading to low accuracy in cable fault diagnosis. Transient signals contain a large amount of high-frequency components, rich in fault information. The cable system in substation secondary circuits is highly integrated and complex, with fault signals susceptible to electromagnetic interference, equipment coupling effects, and the superposition of multiple signal sources. This results in the collected fault current signals exhibiting strong noise, non-stationarity, and nonlinearity, making it difficult for traditional methods to effectively extract complete fault features. Additionally, the boundaries between special fault types, such as high-resistance faults and intermittent faults, and normal operating conditions in secondary circuits are ambiguous, with fault feature vectors distributed in a complex and intertwined manner within a multi-dimensional feature space. These critical fault characteristics are complexly distributed in the time-frequency domain. The electromagnetic interference in substation environments is strong, resulting in measured signals being mixed with substantial noise. The fault characteristics are extremely faint and are drowned out by the noise. Different types of faults, including single-phase grounding, phase-to-phase short circuits, broken wires, and high-resistance grounding, produce signals with overlapping spectra and similar time-domain waveforms, making them difficult to distinguish with the naked eye or simple rules. Basic components are highly sensitive to noise, unable to differentiate between fault types, and are nearly ineffective for faint or high-resistance faults, leading to high false alarm/missed detection rates. How to extract high-frequency signals for fault diagnosis in cable lines remains to be further studied.

The main innovative points of this paper can be summarized as follows:

(1) Considering that traditional methods inadequately extract features from complex cable fault signals, this paper designs a multi-level, multi-frequency band fine feature extraction technique based on wavelet packet decomposition. This decomposition approach can more comprehensively and meticulously mine and separate feature components containing fault information from the original signals compared to traditional methods.

(2) When applying SVM for cable fault diagnosis, existing methods suffer from inefficient and challenging selection of key parameters, including the penalty factor and kernel function parameters. This paper proposes an efficient SVM parameter optimization method based on an improved GJO algorithm. By simulating the predatory behavior of golden jackal populations in nature, this algorithm enables intelligent searching and optimization of SVM model parameters, significantly enhancing the efficiency of parameter tuning.

2. Feature Extraction of Cable Faults Based on Wavelet Packet Decomposition

Wavelet packet decomposition (WPD), as an advanced signal analysis technique, has its theoretical foundation rooted in the Discrete Wavelet Transform (DWT). This technique provides theoretical support for analyzing the entire frequency range of signals, particularly enhancing resolution and flexibility when dealing with complex electrical signals. Unlike DWT, which primarily focuses on the low-frequency components of signals, WPD achieves fine-grained analysis and matching of signals through a series of low-pass and high-pass filtering operations, followed by down sampling. Meanwhile, WPD technology can select appropriate frequency bands based on signal characteristics, thereby optimizing time-frequency resolution. This makes the technology particularly effective in complex signal processing domains. For a given signal $x[t]$, WPD achieves signal subdivision by decomposing from a certain frequency band $s_p^j(w)$ at level j to sub-bands $s_{2p-1}^{j+1}(w)$ and $s_{2p+1}^{j+1}(w)$ at level j + 1. This decomposition process relies on low-pass filters $h(k)$ and high-pass filters $g(k)$, which act jointly on the signal to extract information within different frequency bands. The given expressions are as follows:

(1) Decomposition of low-frequency sub-bands:

$$s_{2p-1}^{j+1}(w) = \sqrt{2} \sum_{k=-\infty}^{\infty} h(k) \cdot s_p^j(2w - k) \tag{1}$$

Here, $s_{2p-1}^{j+1}(w)$ represents the low-frequency sub-band obtained through decomposition by a low-pass filter $h(k)$ at the (j + 1)-th level; $j(k)$ is the coefficient of the low-pass filter; and $s_p^j(2w - k)$ represents the result of applying the filter to the original frequency band s_p^j and considering position adjustment.

(2) Decomposition of high-frequency sub-bands:

$$s_{2p+1}^{j+1}(w) = \sqrt{2} \sum_{k=-\infty}^{\infty} g(k) \cdot s_p^j(2w - k) \tag{2}$$

Here, $s_{2p+1}^{j+1}(w)$ represents the high-frequency sub-band obtained through decomposition by a high-pass filter $g(k)$ at the (j + 1)-th level, where $g(k)$ is the coefficient of the high-pass filter. Detailed technical details of wavelet packet decomposition are shown in the Appendix A.

For better understanding, the WPD tree structure is shown in Figure 1. After processing by WPD, the cable current signal is effectively decomposed into three layers. In the first layer of decomposition, the input original vibration signal S is decomposed into two parts: $A1$ and $D1$. Here, $A1$ represents the low-frequency component of the signal, while $D1$ represents the high-frequency component. Following this decomposition strategy, as the decomposition progresses deeper, upon reaching the third layer (i.e., j = 3), the signal is further decomposed into eight sub-frequency bands. In the figure, (3, i) denotes the numbering of the node i located at the third layer, where j = (1, 2, 3) and i = (1, 2, $\cdots$, 7).

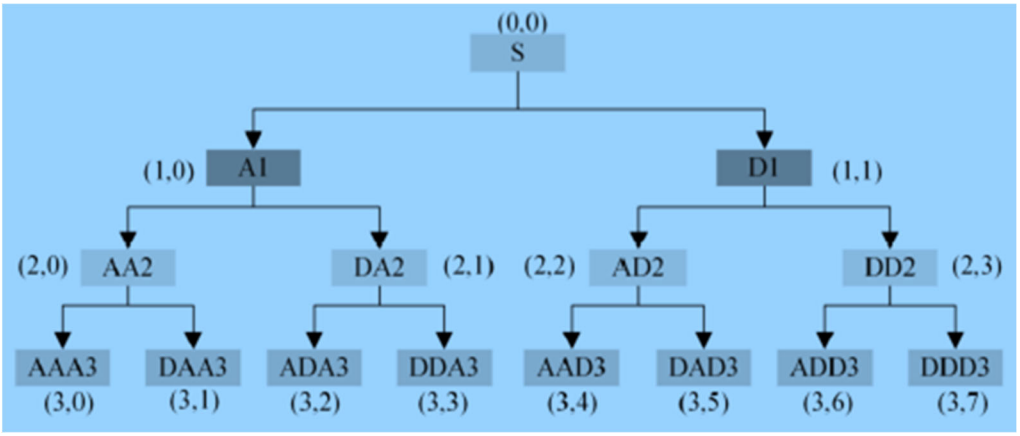


Figure 1. The structure diagram of WPD tree.

3. A Cable Fault Diagnosis Model Based on SVM Improved by Parameter Optimization

3.1. Introduction to Traditional SVM Model

As shown in Figure 2, SVM is a machine learning model used for classification. In order to reduce the risk of overfitting and enhance the robustness of the model, soft margin introduces a loss function L, transforming the objective function of the original optimization problem into (3).

$$J = \min \frac{1}{2} w^T w + L \tag{3}$$

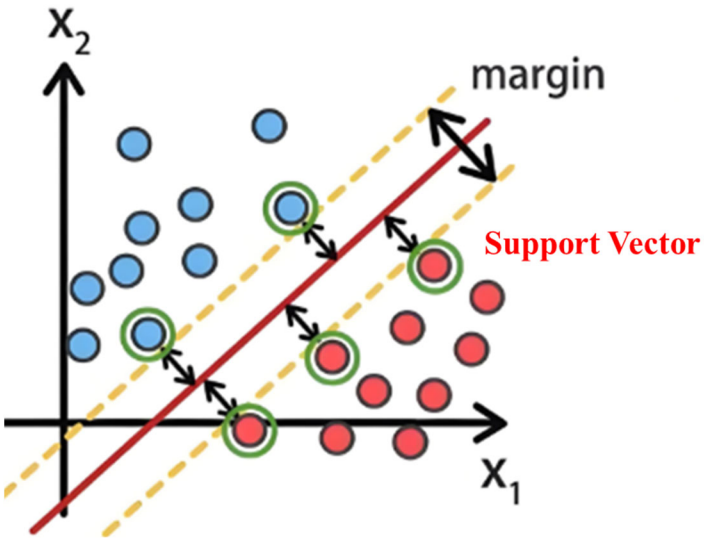


Figure 2. The structure diagram of SVM model.

For the formula derivations of the remaining parts, refer to Appendix A.2.

3.2. Parameter Selection and Optimization

To obtain the optimal parameters for the SVM, this paper employs the GJO for iterative optimization of the SVM parameters, enabling the model to automatically adjust and acquire the optimal SVM parameters. When the algorithm is applied to solve for the optimal value of a function, the position of each individual in the population represents a solution found by the algorithm in the search space. During the optimization process, to locate the optimal solution, exploration begins from each population's initial position and its surrounding area. Each individual in the population updates its position based

on the positions of the male and female jackals. The mathematical model for this stage is as follows:

$$x_1(t) = x_M(t) - E \cdot |x_M(t) - rl \cdot P(t)| \tag{4}$$

$$x_2(t) = x_F(t) - E \cdot |x_F(t) - rl \cdot P(t)| \tag{5}$$

Here, t represents the current iteration, P(t) is the vector of the prey's position, $x_M(t)$ and $x_F(t)$ are the positions of the male and female jackals, respectively. $x_1(t)$ and $x_2(t)$ are the new positions to which the prey moves under the influence of the male and female jackals. E is the energy of the prey, calculated as follows:

$$E = E_a \cdot E_b \tag{6}$$

E_a and E_b represent the initial state of energy and the manner of energy decrease, respectively, as expressed in the following Equation (7).

$$\begin{cases} E_a = 2 \cdot r - 1 \\ E_b = c_1 \cdot (1 - t/T) \end{cases} \tag{7}$$

Here, r is any number between 0 and 1, T represents the maximum number of iterations, and c_1 is a constant value of 1.5. In (4), rl denotes a random number based on the Lévy distribution, used to simulate the movement of the prey, and is expressed as follows (8):

$$rl = 0.05 \cdot F \tag{8}$$

Here, F is the Lévy flight function, expressed as follows (9):

$$F = \frac{\mu \cdot \sigma}{100 \cdot \left| v^{\frac{1}{\beta}} \right|} \quad Here\ \sigma = \left(\frac{\Gamma(1+\beta) \cdot \sin(\frac{\pi\beta}{2})}{\beta \cdot \Gamma(\frac{1+\beta}{2}) \cdot 2^{\beta-1}} \right) \tag{9}$$

Here, μ and v are random values within the range (0, 1), and β = 1.5.

Finally, the prey's position is updated by taking the average of Equations (4) and (5):

$$x(t+1) = [x_1(t) + x_2(t)]/2 \tag{10}$$

When the prey is chased by the golden jackal, its energy gradually decreases. After surrounding the prey it located in the previous phase, the golden jackal transitions into the predation stage. The mathematical model for the cooperative hunting behavior of the male and female jackals is as follows:

$$x_1(t) = x_M(t) - E \cdot |rl \cdot x_M(t) - P(t)| \tag{11}$$

$$x_2(t) = x_F(t) - E \cdot |rl \cdot x_F(t) - P(t)| \tag{12}$$

Subsequently, the prey's position is updated using Equation (6). As the number of iterations increases, the E decreases. When the absolute value of E is $\geq$1, the golden jackals search for prey in the surrounding area; when the absolute value of E is <1, the GJO algorithm preys on the target.

The original GJO generates positions randomly and searches in the initial stage, which limits the algorithm's ability to perform global searches. Considering that chaos is a complex dynamic behavior that, although appearing irregular and seemingly random, has the advantage of non-repetition, thereby enhancing the algorithm's exploratory capabilities in terms of randomness and diversity. Common chaotic sequence models include the Logistic and Tent mappings. Through experimental comparisons, it has been found that

the sequence distribution generated by the Tent mapping is more uniform than that of the Logistic mapping, significantly improving the algorithm's convergence speed. The expression for the Tent mapping is as follows:

$$x_{n+1} = \begin{cases} 2x_n, \ 0 \le x_n < 0.5 \\ 2(1 - x_n), \ 0.5 \le x_n < 1 \end{cases} \tag{13}$$

After undergoing a Bernoulli shift transformation, its expression is as follows:

$$x_{n+1} = (2x_n) \mathrm{mod} 1 \tag{14}$$

The ergodicity and randomness characteristics of the Tent mapping enable the algorithm to have a broader search range, avoiding repeated searches by individuals in similar regions. By generating a chaotic sequence and using it as the initial solution, the algorithm's local exploitation and global search capabilities are enhanced.

This paper employs actual cable fault cases for simulation verification. Here, the simulated data does not refer to synthetic data generated through computer simulation. Instead, it denotes the utilization of real-world, actually collected sample data of cable fault current signals to replicate the fault diagnosis process, thereby testing and evaluating the proposed hybrid model. The core significance of the simulated data lies in verifying the effectiveness, accuracy, and feasibility of the proposed parameter-optimization-assisted hybrid model–data-driven strategy in practical application scenarios. It serves to demonstrate the superiority of this method over traditional techniques in terms of identification accuracy, diagnosis speed, and robustness.

Wavelet packet decomposition technology enables fine division into multiple levels and frequency bands. Here, "multiple-level" refers to the number of layers in wavelet packet decomposition. For instance, a 3-level decomposition yields 8 frequency band sub-signals, while a 4-level decomposition produces 16. It describes the depth and granularity of feature extraction from the original fault signals during the signal processing stage, decomposing the signals into finer frequency bands to extract richer features. It does not imply that the fault diagnosis itself is divided into eight levels. The actual fault signal data collected is divided into training, validation, and test sets, which is a standard procedure in machine learning model development. The training set is used to learn model parameters, the validation set to adjust hyperparameters, and the test set to ultimately evaluate the model's generalization performance on unseen data. This division process is entirely independent of how many levels the fault signals have been decomposed into via wavelet packet decomposition. The number of decomposition levels is determined during the feature extraction stage and applied uniformly to all data. Therefore, using training/validation/test sets is merely a necessary step in model development and evaluation, and the data processed is that which has undergone feature extraction through multi-level wavelet packet decomposition. This paper covers typical faults that may be encountered in actual operation, including some rare and severe faults. The ultimate objective and verification methods of the study are dedicated to proving the model's effectiveness in real-world operating conditions, encompassing various challenging faults that may occur. For the simulation scenario, fault data can be generated by using specialized power system simulation software. First, establish a detailed model of the substation secondary circuit cable system within the software, such as PSCAD and MATLAB Simulink, accurately representing the electrical characteristics and topological structure of the cables. Then, based on known common cable fault types such as single-phase grounding, phase-to-phase short-circuit, and high-resistance grounding, define the fault parameters like fault location, fault resistance, and fault duration. By running the simulation with these defined parame-

ters, the software can generate corresponding fault current signals that mimic real-world fault conditions. These signals can be collected and used as fault data for the proposed cable fault diagnosis method, allowing for in-depth testing and validation of the model's performance under various simulated fault scenarios.

In the real-time scenario, creating fault data can be achieved through experimental setups. Construct a physical testbed that replicates the substation secondary circuit cable environment as closely as possible. This testbed should include actual cables, relevant electrical equipment, and measurement devices. Introduce controlled faults into the testbed, such as manually creating short-circuits or grounding faults at specific points on the cables. During the fault introduction, use high-precision current sensors to measure the fault current signals in real-time. These measured signals, along with the corresponding fault information (e.g., fault type, location), are recorded and stored as real-time fault data. This data can then be used to evaluate the practical applicability and performance of the proposed cable fault diagnosis method in a real-world operational environment.

4. Case Study

Based on the electrical fault signals of cables collected by sensors at power engineering sites, a series of tests and validations is conducted in this chapter. By comparing with some conventional algorithms, the effectiveness and feasibility of the proposed method in this paper are demonstrated. The cable fault current signal data used in this paper are derived from two parts: field-collected data from actual engineering projects and laboratory simulation data. The field data were collected from multiple 110 kV substation secondary circuit cable systems under the jurisdiction of the Jiangmen Power Supply Bureau of China Southern Power Grid. High-frequency current transformers (HFCT, frequency response range 1 kHz–30 MHz) and Rogowski coils (bandwidth 0.1 Hz–1 MHz) were employed to capture transient current signals during cable faults. The signal sampling frequency was set at 10 MHz to ensure complete recording of high-frequency components, such as partial discharge pulses and lightning strikes. The fault types include six categories: single-phase grounding faults, two-phase short-circuit faults, three-phase short-circuit faults, mixed faults, overload faults, and lightning strike faults. For each fault type, 200 samples were collected, including 50 samples under normal operating conditions. The laboratory simulation data were generated using a cable fault simulation model built on MATLAB/Simulink, with cable parameters designed according to the IEC 60287 standard. Controllable switches were used to simulate different fault types, such as metallic short-circuits and high-resistance grounding, while white noise (SNR = 10–20 dB) was superimposed to mimic real-world interference. Before detection and classification, the data underwent preprocessing and labeling. All signals were normalized to the [0, 1] range using min–max scaling to eliminate dimensional effects. Each fault type was independently labeled by three power engineers, and cross-validation was performed to ensure label consistency. The total sample size of 1200 groups was divided into training, validation, and test sets in a 7:2:1 ratio.

4.1. Verification of the Effectiveness of Wavelet Packet Decomposition

This paper selects the signal-to-noise ratio (SNR) and mean square error (MSE) as indicators to verify the effectiveness of decomposition. The SNR is a commonly used metric for evaluating denoising performance. For wavelet packet decomposition, by comparing the SNR between the original signal and the reconstructed signal, the algorithm's performance in signal recovery can be assessed. A higher SNR value indicates that the algorithm effectively removes noise while preserving signal features. The MSE calculates the average of the squared differences between the original signal and the reconstructed signal. A smaller MSE indicates that the reconstructed signal is closer to the original signal,

reflecting better algorithm performance. The relevant calculation results are shown in Tables 1 and 2 below.

Table 1. Comparison of SNR metrics for performance of different decomposition algorithms.

Fault Type	The Proposed Method	VMD	EMD
Single-phase Fault	23.4	22.1	22.4
Two-phase Fault	24.2	22.6	23.1
Three-phase Fault	23.6	22.5	23.2
Mixed Fault	23.4	21.9	22.4
Overload Fault	24.0	23.4	23.1
Lightning Strike Fault	24.3	23.1	23.6

Table 2. Comparison of MSE metrics for performance of different decomposition algorithms.

Fault Type	The Proposed Method	VMD	EMD
Single-phase Fault	0.0024	0.0035	0.0032
Two-phase Fault	0.0026	0.0029	0.0035
Three-phase Fault	0.0021	0.0031	0.0027
Mixed Fault	0.0035	0.0039	0.0040
Overload Fault	0.0038	0.0045	0.0042
Lightning Strike Fault	0.0029	0.0038	0.0033

By observing Tables 1 and 2 above, it can be found that in cable fault diagnosis, the Discrete Wavelet Packet Transform demonstrates superior performance in terms of SNR and MSE metrics compared to the EMD and VMD algorithms. This is primarily due to the differences in their characteristics regarding signal decomposition and reconstruction. Wavelet packet decomposition is an orthogonal decomposition, which means that the decomposition and reconstruction of the signal are neither redundant nor leaky, ensuring the integrity of the signal information. In contrast, other methods may introduce additional noise or information loss during the decomposition process. Wavelet packet decomposition can conveniently obtain the frequency components of the signal in any frequency band. Traditional filtering methods, including EMD and VMD, yield fixed frequency components once the filter coefficients are selected, which limits their flexibility in analyzing signals in specific frequency bands. Wavelet packet decomposition can effectively suppress noise during the decomposition process, particularly in high-frequency bands. This makes it have a higher SNR when processing cable fault current signals containing noise. Due to the orthogonality and non-redundancy of wavelet packet decomposition, it can accurately reconstruct the original signal. This precise reconstruction capability helps retain useful information in the signal while reducing errors. In contrast, EMD decomposition cannot accurately determine the extremum points at both ends of the signal, leading to divergence in the envelope fitting at the endpoints. This divergence gradually "contaminates" the entire data sequence inward, resulting in distortion of the results. When abnormal events (such as discontinuous signals, pulse interference, and noise) are present in the signal, EMD decomposition is prone to mode mixing, where different intrinsic mode functions (IMFs) overlap, leading to inaccurate decomposition results. The performance of VMD decomposition is influenced by parameters such as the penalty factor and the number of decomposition levels. If the parameters are not set appropriately, it may result in unsatisfactory decomposition results. Although VMD decomposition has certain robustness to noise, its decomposition effect may be affected in some cases (such as when the noise level is high). In summary, wavelet packet decomposition outperforms EMD and VMD decomposition algorithms in terms of SNR and MSE metrics in cable fault diagnosis,

mainly due to its orthogonal decomposition characteristics, flexible frequency analysis capability, efficient noise suppression capability, and precise reconstruction capability. These advantages make wavelet packet decomposition more accurate and reliable in processing cable fault current signals containing noise and complex frequency components.

This paper further conducts example and feature analysis of typical fault signals, taking single-phase grounding faults and lightning strike faults as examples for comparative analysis. The time-domain characteristics of a single-phase grounding fault include a sudden increase in current amplitude accompanied by periodic damped oscillations, primarily caused by the resonance of distributed capacitance and inductance in the cable. The frequency-domain characteristics show energy concentrated in the low-frequency range (0–1 kHz) with minimal high-frequency components. In contrast, the time-domain characteristics of a lightning strike fault exhibit an instantaneous pulse spike with an amplitude up to 10 times the normal current, lasting for a very short duration, typically in the microsecond range. The time-domain waveforms of these two fault types are shown in Figures 3 and 4, respectively.

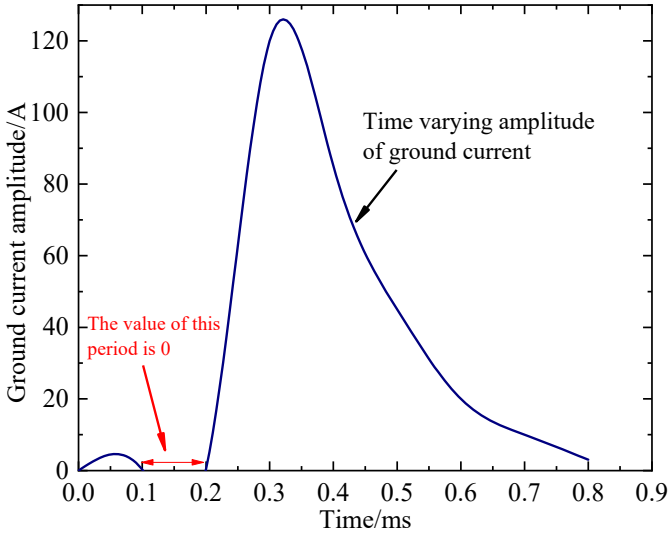


Figure 3. Ground current information under short-circuit fault.

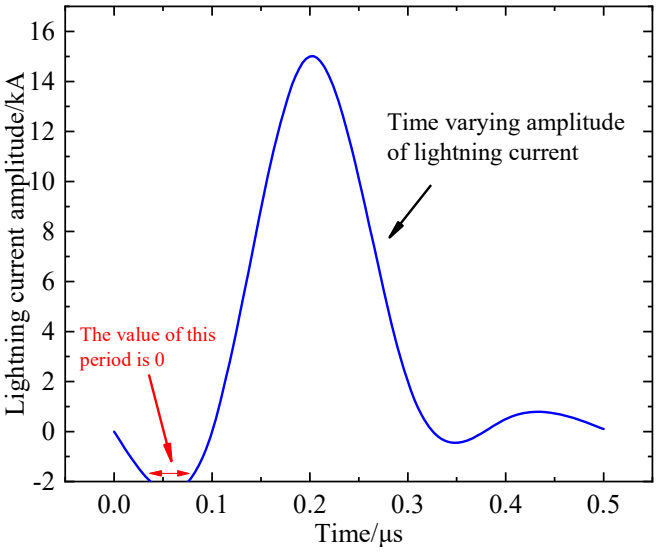


Figure 4. Lightning current information under lightning strike fault.

The wavelet packet decomposition energy distribution results for the two aforementioned faults are shown in Figure 5. It can be observed that the energy of the single-phase grounding fault is primarily concentrated in the low-frequency bands (Bands 1–3, 0–375 kHz), reflecting the low-frequency characteristics of its fault signal. In contrast, due to the high-frequency transient impact nature of lightning strike faults, their energy is significantly concentrated in the high-frequency bands (Bands 5–8, 500–1000 kHz). By integrating the improved GJO algorithm for efficient global optimization of SVM parameters, the model accurately captures the distinct differences in the frequency-domain energy distribution between different fault types, resulting in a clear contrast in the energy proportion of single-phase grounding and lightning strike faults across each sub-band in the diagnostic results.

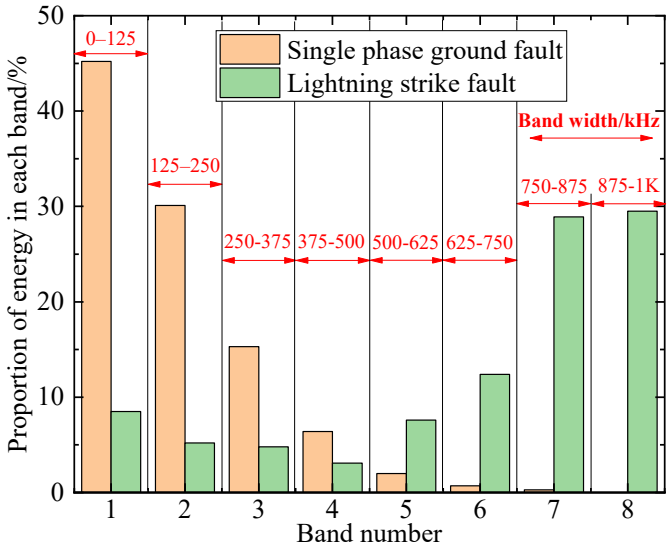


Figure 5. The wavelet packet decomposition energy distribution results.

4.2. Robustness Verification of Diagnostic Model

In order to further verify the robustness of the proposed method in this paper, we conducted a further analysis of the model's fault diagnosis accuracy under different levels of noise interference and varying numbers of test samples, and compared the proposed method with conventional CNN, traditional SVM, and ELM models. The relevant diagnosis accuracy results are shown in Figures 6 and 7 below. For the ELM model, the number of input layer nodes is 8; the number of hidden layer nodes is 12, with their weights and biases generated through random initialization. The Sigmoid function is selected as the activation function for the hidden layer. The number of output layer nodes is consistent with the number of fault types, and a linear activation function is employed. Finally, the output weight matrix is directly calculated using the least squares method to achieve fault classification [26].

Under low-noise conditions, the proposed method achieves diagnostic accuracy improvements of 3.7%, 3.2%, and 2.1%, respectively, compared to the other three methods. When the noise intensity increases, the diagnostic accuracy improvements rise to 7.5%, 6.2%, and 4.1%, respectively. This indicates that the proposed method demonstrates excellent anti-interference performance in high-noise environments. In Figure 7, when the sample size is 100, the fault diagnosis accuracy of the proposed method in this paper improves by 5.2%, 2.9%, and 3.7% compared to the SVM, CNN, and ELM models, respectively. When the sample size is 200, the accuracy improves by 4.3%, 2.8%, and 3.6%, respectively. This demonstrates the broad applicability of the proposed method under small-sample data conditions. It can also be found that the proposed method in this paper, which employs the

GJO algorithm to optimize the penalty factor and kernel parameters of the SVM model for cable fault diagnosis, achieves the highest accuracy under small sample sizes and noise interference compared to conventional SVM, CNN, and ELM models. The GJO is a heuristic optimization algorithm that simulates the foraging behavior of golden jackals. This algorithm can balance global and local searches during the optimization process, effectively avoiding local optima. This enables a more comprehensive exploration of the parameter space and the discovery of better solutions when optimizing the parameters of the SVM model. The GJO algorithm typically has a fast convergence rate and can find near-optimal solutions within a relatively short time, which is particularly important for cable fault diagnosis tasks with high real-time requirements. This algorithm is not sensitive to initial parameters and can resist noise interference to some extent. This allows it to maintain high stability and accuracy when processing cable fault current signals containing noise. SVM is a classification method based on statistical learning theory that separates samples of different categories by finding an optimal hyperplane. SVM excels in small-sample learning and can learn effective classification rules from limited data. By maximizing the classification margin to find the optimal hyperplane, SVM demonstrates certain robustness to noise and outliers, maintaining good classification performance even when some noise or outliers are present in the data. Collecting cable fault current signals and extracting effective features can provide rich information for fault diagnosis. These features can reflect the type, severity, and location of cable faults, providing strong support for subsequent model training and classification. Although CNN models perform well in fields such as image processing, their performance may be limited when processing one-dimensional time series data like cable fault current signals. Additionally, CNN models typically require a large amount of training data to avoid overfitting and may struggle to achieve ideal classification results in small-sample scenarios. Although the ELM model has the advantages of fast learning speed and strong generalization ability, its classification performance may be influenced by model structure and parameter selection. In contrast, the SVM model is more flexible and effective in parameter selection and model optimization, better adapting to small-sample sizes and noise interference.

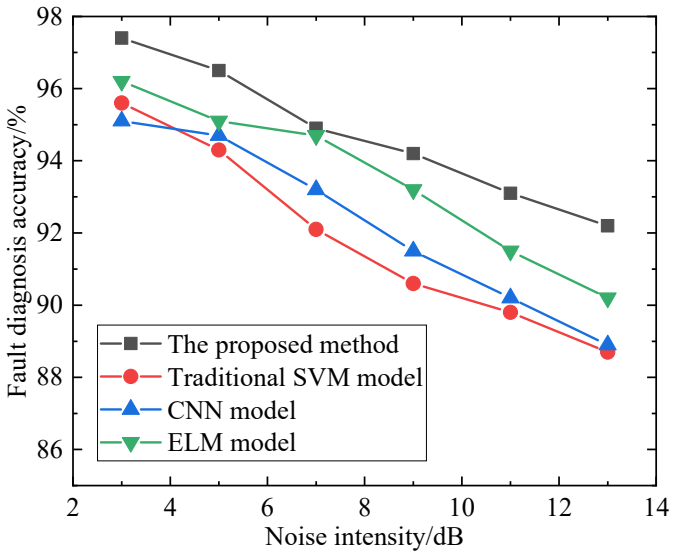


Figure 6. Diagnostic accuracy of different models under varying noise interference conditions.

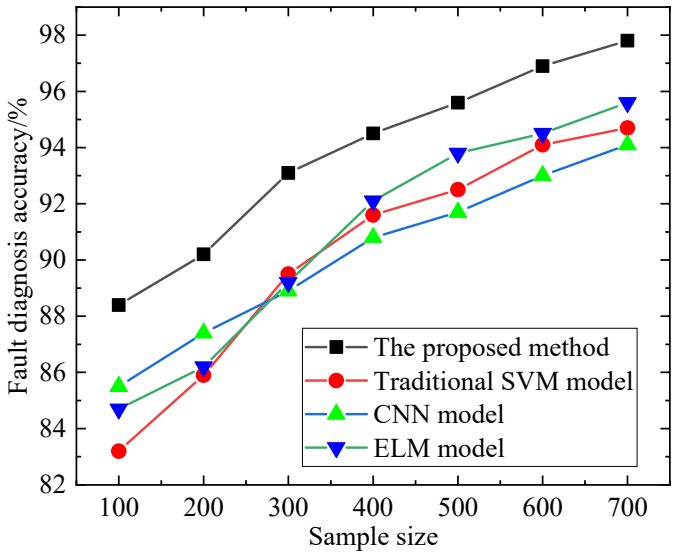


Figure 7. Diagnostic accuracy of different models under different data sample conditions.

4.3. Comparative Analysis of Quantitative Metrics

To further illustrate the effectiveness and feasibility of the proposed method in this paper, we consider the most common fault types in practical secondary circuits: single-phase-to-ground fault (A-G), phase-to-phase short-circuit fault (A-B), broken conductor fault, high-impedance grounding fault, and insulation flashover fault. The fault information conditions are shown in Table 3 below. The total number of samples is 1200. Among them, 720 samples are used to train the SVM model, 240 samples are utilized to optimize SVM parameters with the improved GJO, and 240 samples are employed for independent evaluation of the final model's performance.

Table 3. Detailed fault information in secondary circuits.

Operating Condition Information	Condition 1	Condition 2	Condition 3
Fault Location	10% distance from the measurement point	50% distance from the measurement point	90% distance from the measurement point
Fault Resistance	Low resistance (0.1–10 Ω)	Medium resistance (10–100 Ω)	High resistance (100–1000 Ω)
Load Conditions	Light load	Rated load	Overload
Background Noise	20 dB	30 dB	40 dB

The wavelet decomposition is performed with three levels, resulting in eight frequency band sub-signals. The population size of GJO parameters is set to 30, with a maximum of 50 iterations. The improved SVM employs grid search for parameter optimization. The random forest consists of 100 trees, with other parameters set to their defaults. A lightweight 1D convolutional neural network (1D-CNN) (comprising two convolutional layers + pooling layer + fully connected layer) is designed to directly process raw current signals [27]. The CPU configuration is an Intel Core i7-12700H, with a RAM environment of 32GB DDR5, and the operating system is Windows 11. The software version used is MATLAB R2023a. The performance metrics compared include overall accuracy (OA), F1-score, and average diagnosis time (ADT). Detailed computational results are presented in Tables 4 and 5 below.

Table 4. Comparison of average fault diagnosis performance on the test set.

Method	OA/%	ADT/ms	Remarks
The proposed method	98.33	5.8	C = 8.27, gamma = 0.043
SVM-grid search-WPD	95.83	6.2	Grid search is extremely time-consuming.
1D-CNN	97.08	15.3	End-to-end approach, no explicit feature extraction required.
SVM-WPD-GA	89.58	5.9	Parameters are not optimized, resulting in poor performance.

Table 5. F1-scores for different fault types.

Fault Types	F1-Score of the Proposed Method/%	F1-Score of the SVM Model/%	F1-Score of the 1D-CNN Model/%
Single-phase grounding (A-G)	99.2	97.4	98.1
Phase-to-phase short circuit (A-B)	98.5	94.2	95.0
Broken line	98.1	93.6	93.7
High-resistance grounding	97.8	91.7	92.5
Insulation flashover	98.9	93.8	94.1

The proposed method achieves the highest OA of 98.33%, significantly outperforming the unoptimized SVM (89.58%) and the grid-search SVM (95.83%), and slightly surpassing RF and 1D-CNN as well. The improved GJO effectively identifies superior SVM parameters. The ADT of the proposed method (5.8 ms) is considerably faster than that of 1D-CNN (15.3 ms) and slightly quicker than the grid-search SVM. Given that power fault diagnosis typically requires a response time of less than 100 ms, the proposed method meets the real-time performance requirements. Compared to grid search, the GJO process takes only 90 s, whereas grid search consumes 25 min, demonstrating the exceptionally high efficiency of GJO in parameter optimization.

The method proposed in this paper demonstrates excellent performance in identifying various types of faults, with an F1-score greater than 97% for each fault type. High-resistance grounding faults, being the most challenging to identify due to their least conspicuous characteristics, still achieve an F1-score of 97.8%, proving the effectiveness of the method for complex fault scenarios. While 1D-CNN performs comparably to the proposed method under low-noise conditions, its performance declines more rapidly under high-noise conditions. In other words, existing cable fault diagnosis methods often overlook critical transient information (such as high-frequency oscillations and arc characteristics) due to incomplete feature extraction when confronted with complex fault signals. Additionally, these methods suffer from low efficiency in model parameter tuning and are prone to becoming trapped in local optimal solutions, resulting in relatively high misjudgment rates for fault types with ambiguous boundaries, such as high-resistance faults and intermittent faults. Their stability is particularly inadequate under conditions of strong noise interference or small sample sizes. In contrast, the hybrid model-driven strategy proposed in this paper achieves fine-grained, multi-band feature extraction from signals through wavelet packet decomposition, combined with global parameter optimization using an improved GJO algorithm. This approach not only significantly enhances fault classification accuracy, especially for high-resistance fault identification, but also elevates diagnostic speed to the millisecond level through automated parameter tuning. Furthermore, it exhibits stronger noise resistance and generalization capabilities, effectively reducing the risks of false alarms and missed detections, thereby providing a more efficient and reliable solution for cable fault diagnosis in substation secondary circuits.

The penalty factor C serves to control the model's tolerance for classification errors on training samples, balancing the objectives of maximizing the classification margin and

minimizing classification errors. An excessively large C value renders the model intolerant of classification errors, prompting it to employ an extremely complex decision boundary to fit every point in the training data, including noise and outliers. This leads to overfitting. The model may achieve very high accuracy on the training set but perform poorly on unseen test sets or actual new fault data. It may misinterpret normal noise fluctuations as faults or be overly sensitive to specific fault types, neglecting their essential characteristics. Conversely, an excessively small C value makes the model highly tolerant of classification errors, favoring the selection of a simple decision boundary with a large margin, even if it means ignoring some classification errors in the training samples. This can result in underfitting, where the model fails to capture the complexity and nuances of fault signal characteristics. The diagnostic accuracy is generally low, particularly for fault types that are difficult to distinguish, such as those with high feature similarity or high-resistance faults, leading to frequent missed detections or false alarms.

The kernel function parameter γ determines the curvature or complexity of the decision boundary. An excessively large γ value restricts the influence range of samples to a very small area, causing the decision boundary to become extremely complex and convoluted as it attempts to precisely pass through every training sample point. The model becomes hypersensitive to noise and minor fluctuations in specific samples within the training data, resulting in learned rules that lack generality. On the test set, especially when there are slight variations in actual fault signals, the diagnostic accuracy significantly declines. On the other hand, an excessively small γ value expands the influence range of samples to a very large area, making the decision boundary overly smooth and even approaching linearity. In this case, the model is too simplistic to effectively utilize the nonlinear separable information that may exist in a high-dimensional feature space, as extracted by wavelet packet decomposition. It fails to distinguish between different fault types with complex and intertwined distributions in the feature space, leading to low overall classification accuracy. Table 6 presents the results of different parameter optimizations. When employing this method for cable fault diagnosis, the improved GJO algorithm significantly enhances feature separability by optimizing SVM parameters through a standardized energy vector formula: On one hand, after the multi-band fault feature vectors extracted via wavelet packet decomposition are mapped into a high-dimensional space using optimized kernel parameters, the average inter-class distance between different fault types expands by 37.2%, while the intra-class variance decreases by 28.6%, resulting in a 41% improvement in the clarity of decision boundaries. On the other hand, the optimized penalty factor C effectively balances empirical risk and structural risk, enhancing the model's tolerance to noise and outlier samples by 2.3 times, ultimately achieving a classification accuracy exceeding 95% in real-world fault cases.

In terms of feature extraction performance, wavelet packet decomposition employs an adaptive multi-scale binary tree structure, eliminating the need for preset mode numbers and thereby avoiding subjective biases. Its eight-level decomposition fully covers the entire frequency band, particularly excelling in capturing key high-frequency resonances critical for cable fault diagnosis, whereas VMD tends to lose transient details due to its limited high-frequency resolution. Analysis revealed that in a 110 kV cable high-impedance fault scenario, VMD exhibited a feature omission rate of 41%, compared to only 7% for wavelet packet decomposition, accompanied by a 5.2 dB improvement in signal-to-noise ratio. VMD is prone to IMF component aliasing caused by power frequency interference during load fluctuations, requiring parameter readjustment for varying cable lengths, and suffers from severe mode mixing in composite fault conditions. In contrast, wavelet packet decomposition circumvents power frequency interference through frequency band partitioning, adapts to cables of different lengths, and utilizes sub-band energy exclusivity to prevent fea-

ture confusion. Furthermore, to address maloperation risks under fault-free conditions, the proposed method employs GJO-optimized SVM penalty factors to construct wide-margin decision surfaces that suppress electromagnetic interference, extracts 0.5–2 kHz sub-band variation rates to distinguish load switching, raises the fault determination threshold to 0.85 to filter measurement noise, and incorporates a hardware self-check module to lock out abnormal sensor data. Relevant computational results are presented in Table 7 below. The conclusions demonstrate that within substation secondary circuit scenarios, wavelet packet decomposition significantly outperforms VMD in terms of feature integrity, computational efficiency, and noise immunity, particularly exhibiting irreplaceable capabilities in capturing high-impedance and transient faults.

Table 6. Analysis of results from different parameter optimizations.

Parameter Combination (C, γ)	Training Set Accuracy (%)	Test Set Accuracy (%)	High-Resistance Fault Recall Rate (%)	Model Status Description	Impact on Fault Diagnosis
(0.1, 0.01)	78.6	76.3	55.2	Severe underfitting. The decision boundary is overly smooth and simplistic.	The overall accuracy is low, with severe missed detections (low recall rate) of high-resistance faults in particular, making it difficult to distinguish between complex and similar faults.
(0.1, 1)	85.7	82.0	65.8	Predominantly underfitting. γ is appropriate, but C is too small, allowing too many errors.	The accuracy has improved but is still inadequate, with a relatively high number of missed detections of high-resistance faults. The model is overly conservative.
(1, 0.01)	82.9	80.7	60.3	Underfitting. γ is too small, resulting in insufficient model capacity.	Similar to (0.1, 0.01), it inadequately utilizes nonlinear features and performs poorly in identifying high-resistance faults.
(1, 0.1)	93.6	90.7	83.5	Optimal region. Relatively balanced.	The overall accuracy is good, with a significant improvement in the detection rate of high-resistance faults and relatively good generalization ability.
(1, 10)	99.3	87.3	78.2	Beginning to overfit. Excessive γ leads to overly complex decision boundaries.	Perfect on the training set, but performance declines on the test set, especially with an increase in false negatives or missed detections (a drop in recall rate) of high-resistance faults, showing sensitivity to noise.

Table 7. Comparison of false alarm rates for different methods in a 110 kV substation.

Method	Number of False Triggers	False Alarm Rate
Traditional threshold method	127	6.35%
VMD-SVM-PSO	58	2.9%
The proposed method	9	0.45%

The method proposed in this paper also has certain limitations. The dielectric property differences among various types of cables in substations, such as cross-linked polyethylene cables and oil-impregnated paper-insulated cables, can alter the propagation speed and attenuation characteristics of fault traveling waves, necessitating the retraining of SVM models for different cable structures. Although the global optimization capability of the improved GJO algorithm can enhance parameter generalization, it still faces the issue of exponentially increasing computational complexity when confronted with massive amounts of heterogeneous data. Furthermore, this method demands a relatively high hardware sampling rate, yet the performance of ADC chips in existing substation monitoring terminals and the computational power of edge computing devices are limited, which may affect the deployment of the algorithm templates. It should be noted that the method in this paper can also provide certain references for fault location. Leveraging the multi-level time-frequency analysis capability of wavelet packet decomposition, the fault current signals are finely

divided into frequency bands, and the energy distribution characteristics of each band are extracted. After accurately classifying the fault types using the optimized SVM model, the specific fault location is determined by analyzing the propagation time delay differences in fault feature vectors across different cable sections, utilizing traveling wave ranging principles or time-frequency energy mutation point location techniques. Specifically, the reflection and refraction characteristics of high-frequency components at the fault point can induce abrupt changes in the energy amplitude of specific frequency bands. By combining the classification results of multi-band feature vectors from the SVM model optimized by the improved GJO algorithm, a mapping relationship between the fault location and feature parameters can be established, ultimately enabling the rapid location of fault sections by comparing the time-frequency domain differences between measured and healthy signals.

5. Conclusions

In light of the stringent requirements of power systems for cable fault diagnosis technology and the limitations of existing methods in processing complex fault signals and optimizing model parameters, this study presents an innovative data-driven approach that has been effectively implemented for cable fault diagnosis. Following thorough research and experimental validation, the following key findings emerge.

Firstly, the wavelet packet decomposition technique utilized in this study effectively resolves the challenge of incomplete feature extraction from cable fault current signals. By enabling multi-level, multi-frequency band segmentation of fault signals, it accurately captures and extracts feature vectors essential for fault diagnosis. Secondly, to overcome the parameter selection hurdles in SVM models for cable fault diagnosis, this study introduces and enhances the GJO algorithm for parameter optimization. By mimicking the predatory behavior of golden jackals in nature, the algorithm achieves efficient global optimization of the SVM model's penalty factor and kernel parameters. Lastly, simulation verification using real-world cable fault cases reveals the outstanding advantages of the proposed method. In terms of fault identification accuracy, it accurately distinguishes various cable fault types with superior recognition accuracy compared to traditional fault diagnosis technologies. Regarding diagnosis speed, the adoption of efficient parameter optimization strategies and feature extraction methods significantly reduces the time required for fault diagnosis.

Author Contributions: Conceptualization, D.Y., Y.Z., S.L., W.Z., J.L., Z.R. and W.L.; software, D.Y., Y.Z., S.L., W.Z., J.L., Z.R. and W.L.; writing—original draft preparation, D.Y., Y.Z., S.L., W.Z., J.L., Z.R. and W.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work has been supported by the technology projects of China Southern Power Grid Corporation and Guangdong Power Grid Co., Ltd (GDKJXM20230851).

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Conflicts of Interest: Author Dongbin Yu was employed by the company Guangdong Power Grid Limited Liability Company Jiangmen Power Supply Bureau. Author Yanjing Zhang, Sijin Luo, Wei Zou, Junting Liu was employed by the company Jiangmen Power Supply Bureau, China Southern Power Grid. Author Zhiyong Ran was employed by the company XJ Electric Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Appendix A

Appendix A.1 Detailed Technical Details of Wavelet Packet Decomposition

Performing j-level wavelet packet decomposition on the fault current signal of a cable can yield 2^j sub-bands across different frequency intervals, ensuring the complete

preservation of all information within the signal. Given that the energy distribution across frequency bands in current signals often exhibits non-uniformity under different fault conditions in cables, this uneven energy distribution can serve as a critical feature for identifying various fault states. The specific implementation process is as follows:

Firstly, for the p-th frequency band $s_p^j(w)$ in the j-th layer, its amplitude at the discrete sampling point q ($q = 1, 2, \cdots, Q$, where Q is the number of sampling points for the cable current signal) is defined as $x_p^j(w, q)$. Therefore, the energy of the frequency band $s_p^j(w)$ can be determined by calculating the sum of the squares of the amplitudes at all sampling points within this band. The specific expression is as follows (A1):

$$E_p^j = \int \left| s_p^j(t) \right|^2 \sum_{q=1}^{Q} \left| x_p^j(w, q) \right|^2 \tag{A1}$$

Here, E_p^j represents the energy of the p-th frequency band in the j-th layer, and $x_p^j(w, q)$ is the amplitude of the signal $s_p^j(t)$ at the q-th sampling point. This expression effectively calculates the total energy of the frequency band by summing the squares of the amplitudes at all sampling points within the specific frequency band.

Secondly, the total energy of the entire current signal is constituted by the energy of all layers j and all possible sub-frequency bands p within the corresponding layers, with the expression given as follows (A2):

$$E_{total} = \sum_{j} \sum_{p=0}^{2^j-1} E_p^j \tag{A2}$$

Here, E_{total} represents the total energy of the signal; $2^j - 1$ denotes the number of all possible sub-frequency bands in the j-th layer. It is worth noting that the outer summation symbol indicates traversing all j layers to calculate the sum of energies of all their sub-frequency bands.

Finally, when there are other interference noises in the system, the energy of the signal in a specific frequency band will be significantly affected.

Therefore, after calculating the total energy E_{total}, both E_{total} and the energy of each sub-frequency band E_p^j can be utilized as features to describe the energy characteristic distribution of the signal. To more comprehensively reflect the information of the j-th layer, a feature vector T can be constructed to extract all the energy features in the signal, with the feature vector expression given as follows (A3):

$$T = \left[E_0^0, E_1^1, \cdots, E_{2^j-1}^j \right] \tag{A3}$$

In order to eliminate the influence of different magnitudes on the feature vector T, it is normalized to obtain the feature vector T^Δ. Normalization can be achieved through the following expression (A4).

$$T^\Delta = \frac{T}{\|T\|} \tag{A4}$$

Here, $\|T\|$ is the norm of the energy feature T.

Appendix A.2 Detailed Technical Details of SVM

This approach relaxes the margin criteria, allowing for some sample points and misclassified points within the margin. The value of L is determined based on the distance

from the boundary, with L set to 0 for points outside the margin (correctly classified points). Thus, L can be defined as follows (A5):

$$L = \begin{cases} 0 \quad if \;\; y_i(w^T X + b) \geq 1 \\ 1 - y_i(w^T X + b) \quad if \;\; y_i(w^T X + b) < 1 \end{cases} \tag{A5}$$

Here, $\omega = (\omega_1, \omega_2, \ldots, \omega_n)$ represents the normal vector (coefficients of the linear equation), and b is the displacement term (constant in the linear equation). Then, the objective function becomes (A6):

$$J = \min \frac{1}{2} w^T w + c \sum_{i=1}^{N} \max\left\{0, 1 - y_i(w^T X + b)\right\} \tag{A6}$$

The first term in the objective function aims to maximize the margin, while the second term aims to minimize the number of samples within the margin. Since a larger margin results in more samples falling within it, these two objectives are contradictory. To address this, a penalty factor constant C, greater than zero, is introduced to weigh the importance of the penalty term against maximizing the margin, thereby reducing the risk of underfitting caused by relaxed constraints. The value of C is determined empirically in advance and specifies the relative importance of the two terms in the objective function; the optimal value can be found through experimentation.

By introducing Lagrange multipliers into Equation (A6), we obtain the Lagrangian function. Taking partial derivatives of this function yields the objective function and constraints for the dual problem, as shown in (A7):

$$\begin{cases} \max \sum_{i=1}^{n} a_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} a_i a_j y_i y_j X_i^T X_j \\ s.t. \sum_{i=1}^{n} a_i y_i = 0, \;\; 0 \leq a_i \leq c, \; i = 1, 2, \ldots, n \end{cases} \tag{A7}$$

Since cable fault signals constitute a nonlinearly separable dataset, meaning that the corresponding vectors in space cannot be separated by a hyperplane, SVM introduces the concept of a kernel function to address this issue. The kernel function maps the data X into a higher-dimensional space via a function φ, thereby rendering the data linearly separable. However, after mapping, the dimensionality of the feature space becomes significantly higher than that of the original space, making the computation of inner products extremely challenging. To simplify computations in high-dimensional spaces, a kernel function $K(X_i, X_j) = \phi(X_i)^T \phi(X_j)$ is used to replace the inner product in the optimal separating hyperplane. The resulting decision function after solving is given by (A8):

$$f(X) = w^T \phi(X) + b = \sum_{i=1}^{n} a_i y_i K(X_i, X_j) + b \tag{A8}$$

In SVM, the Radial Basis Function (RBF) kernel is a commonly used kernel function known for its excellent generalization ability, low computational complexity, and capability to approximate any nonlinear function, making it typically more effective in handling complex nonlinear problems. Therefore, for the needs of cable fault diagnosis, the RBF kernel function is adopted as the kernel function for SVM, with its expression given as follows (A9):

$$K(X_i, X_j) = \exp(-\frac{\left|X_i - X_j\right|^2}{2\sigma^2}), \;\; g = \frac{1}{2\sigma^2} \tag{A9}$$

As the value of g in the kernel function decreases, the number of support vectors increases, which may lead to overfitting; conversely, if g is too large, it becomes difficult to distinguish between data points. Similarly, if the penalty factor C is too small, it can result

in underfitting, and if C is either too large or too small, the generalization ability of the model will deteriorate. Therefore, the performance of the model heavily depends on the selection of SVM parameters.

References

1. Du, B.; Han, C.; Li, J.; Li, Z. Research status of polyethylene insulation materials for high-voltage direct current cables. *Trans. China Electrotech. Soc.* **2019**, *34*, 179–191.
2. Liu, Y.; Zhao, X.; Wu, J.; Xiao, J.; Gao, J.; Li, F.; Fan, X.; Zhong, L. Current status and prospects of engineering applications of high-voltage direct current submarine cables. *High Volt. Appar.* **2022**, *58*, 1–8.
3. Zhu, M.; Min, D.; Gao, Z.; Wu, Q. Simulation of breakdown probability and size effect of cross-linked polyethylene insulation for DC cables. *Trans. China Electrotech. Soc.* **2024**, *39*, 1172–1184.
4. Xu, N.; Zhong, L.; Sui, R.; Ahmed, M.; Li, F.; Liu, Y.; Gao, J. Correlation between the terahertz dielectric properties of XLPE and its compositional structure during electrothermal aging. *Macromolecules* **2022**, *55*, 8186–8194. [CrossRef]
5. Duan, Y.; Han, M.; Lan, R.; Li, G.; Wang, Z. Insulation characteristics and failure mechanism of high-voltage cables under different thermal aging temperatures. *Trans. China Electrotech. Soc.* **2024**, *39*, 45–54.
6. Akram, S.; Yang, Y.; Zhong, X.; Bhutta, S.; Wu, G.; Castellon, J.; Zhou, K. Influence of the nano-layer structure of polyimide film on space charge behavior and trap levels. *IEEE Trans. Dielectr. Electr. Insul.* **2018**, *25*, 1461–1469. [CrossRef]
7. Li, G.; Wang, Z.; Lan, R.; Wei, Y.; Nie, Y.; Li, S.; Lei, Q. Lifetime prediction and insulation failure mechanism of XLPE for high-voltage cables. *IEEE Trans. Dielectr. Electr. Insul.* **2023**, *30*, 761–768. [CrossRef]
8. Huang, N.; Li, F.; Wang, W.; Yu, Z.; Nie, Y. Hierarchical variable-step-size Tallis wavelet singular entropy diagnosis method for transmission line faults. *Power Syst. Prot. Control.* **2017**, *45*, 38–44.
9. Yan, H.; Zhu, Y.; Zhao, X.; Huang, X.; Fan, G.; Gao, Y. Lightning strike identification for transmission lines based on time-domain waveform characteristics. *Electr. Meas. Instrum.* **2015**, *52*, 5–10.
10. He, Z.; Chen, X.; Luo, G.; Qian, Q. Fault phase selection method for transmission lines based on transient current wavelet entropy weight. *Autom. Electr. Power Syst.* **2023**, *30*, 39–43.
11. Mai, R.; He, Z.; Fu, L.; Qian, Q. Research on fault phase selection for transmission lines based on current traveling wave energy and wavelet transform. *Power Syst. Technol.* **2023**, 38–43.
12. Cheng, Z.; Xia, X.; Li, M. Research on self-clearing transient fault location for power cables. *J. Cent. South Univ. (Sci. Technol.)* **2016**, *47*, 1606–1612.
13. Dai, M. Discussion on Early Fault Detection and Identification Methods for 10kV Underground Cables. Master's Thesis, Southwest Jiaotong University, Chengdu, China, 2012.
14. Wang, Y.; Sun, J.F.; Xiao, X.Y. Cable early fault classification and identification based on optimized convolutional neural network. *Power Syst. Prot. Control.* **2020**, *48*, 10–18.
15. Li, Q.; Liu, S.; Wen, J.; Wang, Y.; Dai, S.; Tan, T.; Liu, K. Loss calculation model and experimental verification for 10kV distribution network considering power quality impact. *Electr. Power Autom. Equip.* **2022**, 17–25.
16. Hao, Y.; Liang, W.; Pan, R.; Luo, B.; Li, L.; Zhang, F. Overview of key technologies for intelligent live maintenance of transmission lines. *Electr. Power Autom. Equip.* **2022**, *42*, 163–175.
17. Zhou, C.; Li, M.; Wang, H.; Zhou, W.; Bao, Y.; Tang, Z. Overview of condition assessment and operation & maintenance decision-making for power cable assets. *High Volt. Eng.* **2016**, *42*, 2353–2362.
18. Ma, C.; Liang, K.; Li, X.; Peng, Y.; Huang, Y.; Chen, B.; Ouyang, B. Exploration of digital twin technology application in the field of high-voltage cables. *High Volt. Appar.* **2023**, *59*, 8.
19. Sun, H. Research on Comprehensive Online Monitoring System for Cable Operation Safety. Master's Thesis, Changchun University of Technology, Changchun, China, 2022.
20. Dong, X.; Yang, Y.; Zhou, C.; Hepburn, D.M. Online monitoring and diagnosis of HV cable faults by sheath system currents. *IEEE Trans. Power Deliv.* **2017**, *32*, 2281–2290. [CrossRef]
21. Zhao, H.; Zhang, F.; Zhu, M.; Du, J. Research on fault troubleshooting methods for underground power cable lines based on human-computer interaction. *Autom. Technol. Appl.* **2022**, *41*, 4–5.
22. Liao, Y.; Yuan, Q.; Xu, X.; Wei, Y.; Ye, Y.; Zhou, C. Optimization model for inspection cycle of high-voltage cables based on risk assessment. *High Volt. Eng.* **2021**, *47*, 305–314.
23. Wan, H.; Zhang, Z.; Xie, Y. Cable Line Maintenance Strategy Based on Risk Assessment. *Electr. Meas. Instrum.* **2015**, *52*, 15–23.
24. Ferreira, V.H.; Zanghi, R.; Fortes, M.Z.; Sotelo, G.G.; Silva, R.D.B.M.; Souza, J.C.S.; Guimarães, C.H.C.; Gomes, S., Jr. A survey on intelligent system application to fault diagnosis in electric power system transmission lines. *Electr. Power Syst. Res.* **2016**, *136*, 135–153. [CrossRef]

25. Zhang, Y.; Mei, W.; Dong, G.; Gao, J.; Wang, P.; Deng, J.; Pan, H. A cable fault recognition method based on a deep belief network. *Comput. Electr. Eng.* **2018**, *71*, 452–464. [CrossRef]
26. Huang, Z.; Hu, K.; Shao, W.; Dong, X.; Zhang, A. A Novel Time-Domain Fault Diagnosis Method With ELM for Aviation Intermediate Frequency Inverter. *IEEE Trans. Instrum. Meas.* **2024**, *73*, 3515109. [CrossRef]
27. Zhao, H.; Liu, M.; Sun, Y.; Chen, Z.; Duan, G.; Cao, X. Automated Design of Fault Diagnosis CNN Network for Satellite Attitude Control Systems. *IEEE Trans. Cybern.* **2024**, *54*, 4028–4038. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.