Lian K, Xie S, Sun F, Luo Y. Analysis of SVM-SMOTE algorithm and Adaboost-RBF algorithm for low-voltage switch fault diagnosis. Diagnostyka. 2026;27(4):2026402. https://doi.org/10.29354/diag/242669.

DIAGNOSTYKA, 2026, Vol. 27, No. 4

e-ISSN 2449-5220
DOI: 10.29354/diag/242669

ANALYSIS OF SVM-SMOTE ALGORITHM AND ADABOOST-RBF ALGORITHM FOR LOW-VOLTAGE SWITCH FAULT DIAGNOSIS

Kaimin LIAN, Sheng XIE *, Feng SUN, Yuanping LUO

Guangdong Power Grid Co., Ltd. Jieyang Jiedong Power Supply Bureau, Jieyang, 522000, China
* Corresponding author, e-mail: XS45698752@163.com

Abstract

Low-voltage switch fault diagnosis is challenged by imbalanced fault samples and insufficient recognition of minority classes. This study proposes a hybrid SVM-SMOTE-AdaBoost-RBF (SSAR) model that combines SVM-guided minority-class oversampling with adaptive ensemble-based nonlinear classification. In the proposed framework, SVM-SMOTE uses decision-boundary information to support minority-class sample balancing, while AdaBoost-RBF adaptively adjusts sample weights and performs fault classification. Experimental results show that SVM-SMOTE achieved a data completeness rate of 96.45% on the CWRU-CB dataset and a fault-type recognition accuracy of 98.11% on the NASA-Relay dataset. On the self-generated dataset, SSAR achieved a diagnostic accuracy of 98.24%, a precision of 97.11%, a recall of 95.98%, and an F1-score of 96.48%, with a diagnosis time of 2.08 s for 100 iterations. These findings indicate that integrating boundary-guided sample balancing with adaptive nonlinear classification can improve fault diagnosis under imbalanced data conditions. Further validation under realistic operating conditions is required to establish its practical applicability.

Keywords: low-voltage switch; fault diagnosis; support vector machine; oversampling technique; adaptive boosting algorithm

List of Abbrevations:

SVM - Support Vector Machine
SMOTE - Synthetic Minority Over-sampling Technique
AdaBoost - Adaptive Boosting
RBF - Radial Basis Function
SSAR - SVM-SMOTE-AdaBoost-RBF
FFT - Fast Fourier Transform
TLs - Tomek Links
EMD - Empirical Mode Decomposition
NBC - Naive Bayes Classifier
DBN - Deep Belief Network
GMM - Gaussian Mixture Model
VMD - Variational Mode Decomposition
MTF - Markov Transition Field
GAN - Graph Attention Network
CNN - Convolutional Neural Network
BKA - Bi-directional Kernel Attention
MAE - Mean Absolute Error
CWRU-CB - Case Western Reserve University Circuit Breaker Data Set
NASA-Relay - NASA Relay Aging and Life Prediction Data Set

1. INTRODUCTION

As core control components maintaining stable operation at the distribution network terminals, low-voltage switches have emerged as a critical focus of global research for research on improving their quality and reliability [1]. Accurate fault diagnosis can prolong component service life, consequently, extensive research has been conducted in this field. Guerrero et al. proposed a detection method for low-voltage switches and electrical faults affected by current. This method distinguished the affected areas by sequentially switching resistors in the system current, measured neutral voltage, and completed fault identification [2]. An adaptive low-voltage ride-through method was proposed by Luo et al. to mitigate inverter damage resulting from low-voltage current faults. This method improved the efficiency of fault detection by compensating voltage differences, connecting voltages, and adjusting output power [3]. Aluri et al. proposed a fault location method based on parameter estimation to accurately detect fault locations in low-voltage switches and power grids. This method introduced natural oscillation frequency calculation and Fourier amplitude spectrum to identify the oscillation frequency of fault paths, improving diagnostic accuracy [4]. Related research on permanent magnet synchronous motors (PMSMs) provides additional context for intelligent electrical-fault diagnosis. Ziad et al. reviewed PMSM fault-detection techniques and classified them into signal-based, model-based, and data-driven approaches [30].

Received 2026-01-22; Accepted 2026-11-09; Available online 2026-11-10

© 2026 by the Authors. Licensee Polish Society of Technical Diagnostics (Warsaw. Poland). This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).

Faraj et al. investigated a deep-learning-based method for classifying electrical faults in PMSMs [31]. In a comparative study, Ziad et al. evaluated deep-learning models using scalogram images derived from stator-current signals and reported 100% accuracy, precision, recall, and F1-score for ResNet50 on their experimental dataset [32]. Although these studies demonstrate the potential of data-driven methods for electrical-fault recognition, they focus on motor-specific fault characteristics and do not directly address imbalanced low-voltage switchgear data through the combined use of boundary-guided oversampling and adaptive nonlinear classification. This gap motivates the development of the proposed SSAR framework. Current deep-learning-based methods for processing fault sample data require further data expansion and balanced to address the class imbalance problem occurring in the diagnostic process, and some methods still have difficulty identifying rare fault features [5]. Support Vector Machine (SVM), as a supervised learning classification algorithm, can accurately classify different data and features through its structure and effectively handle high-dimensional data [6,7]. Ali et al. proposed a diagnostic system based on SVM to improve the accuracy of partial discharge detection. This system extracted features through SVM, then refined and classified features using deep neural networks to complete detection [8]. Nzel H et al. proposed a hybrid model combining artificial neural networks and random forests, employing the Optuna framework for automatic hyperparameter optimization, which demonstrated excellent performance in transmission system fault detection. These findings further validate the feasibility of employing hybrid models and intelligent optimization methods for power equipment diagnosis [9]. The Synthetic Minority Over-sampling Technique (SMOTE) addresses class imbalance by generating synthetic minority-class samples [10,11]. The team led by Nugroho proposed a technique combining extreme gradient boosting with SMOTE to classify diabetes. This technique combined bidirectional long short-term memory and bidirectional gated recurrent units with extreme gradient boosting, and solved data imbalance through SMOTE to complete classification. Experimental results showed that the F1 score of this technique reached 100% [12]. In addition to data-driven approaches, model-based diagnostic frameworks have also been applied in complex systems. Scholars such as Guemmour M. B. combined fault tree analysis with Bayesian networks for fault diagnosis in wave energy conversion systems, demonstrating the effectiveness of hybrid reasoning models in uncertain environments [13]. Therefore, the SVM and SMOTE algorithms are integrated to develop the SVM-SMOTE framework, thereby enhancing data analysis capabilities in fault diagnosis.

With fault diagnosis data collection as the foundation, the efficiency of fault identification and analysis is further improved. However, traditional low-voltage switch fault detection still has problems of susceptibility to noise influence and insufficient generalization performance and diagnostic accuracy. The Adaptive Boosting (AdaBoost) algorithm enhances feature recognition precision and adaptability by integrating multiple classifiers, facilitating automated feature selection [14,15]. Li et al. proposed an algorithm integrating AdaBoost, gradient boosting decision tree, and extreme gradient boosting to improve the prediction accuracy of mineral resource discovery. This algorithm used SMOTE to generate training SMOTE-AdaBoost-RBF composite model for low-voltage switch fault diagnosis, abbreviated as SSAR model, which is expected to solve the class imbalance problem in fault diagnosis, improve diagnostic accuracy and efficiency, ensure stable operation of low-voltage switches, and thereby enhance the maintenance and operation capability of the entire distribution network. Distinct from traditional simple concatenation models, this approach enables end-to-end high-precision fault diagnosis for low-voltage switches under imbalanced data conditions. datasets, and then drew mineral resource prospectivity maps through AdaBoost and other algorithms [16]. In the field of wind turbine fault diagnosis, Li M et al. proposed a data-driven approach based on residual analysis, which achieves efficient fault isolation by constructing residual signals and setting adaptive thresholds. The study demonstrates that integrating signal processing with data-driven strategies significantly enhances the robustness of diagnostic systems [17]. Radial Basis Function (RBF) networks are also used for nonlinear modeling and function approximation [18,19]. Zhang et al. proposed an interpolation algorithm based on RBF to reconstruct radiation fields. This algorithm reconstructed source radiation fields through RBF, then used probabilistic roadmap to plan paths, improving operator safety [20]. Therefore, this study combines AdaBoost with RBF algorithm to obtain the AdaBoost-RBF algorithm for diagnosing low-voltage switch fault problems. Finally, this study integrates the SVM-SMOTE algorithm and AdaBoost-RBF algorithm to obtain the final SVM- To further clarify the distinction between the proposed mechanism and conventional SMOTE or a simple AdaBoost-RBF combination, the proposed framework can be formulated mathematically. First, the SVM decision function $f(x)$ is used to identify minority samples close to the classification boundary, and the difficult-sample set is defined as $B = \{x_i : |f(x_i)| \leq \tau\}$, where τ denotes the boundary threshold. SMOTE interpolation is then performed using samples in B, constraining the generated samples according to the SVM decision boundary. For the subsequent diagnostic module, the sample weights $\omega_i^{(k)}$ generated by AdaBoost are introduced into the weighted training of the RBF network, assigning greater optimization weights to misclassified samples. Therefore, SSAR establishes a continuous mechanism of " decision-boundary -constrained sampling – misclassification

Table 1. Quantitative comparison of representative studies and the research gap

Study	Method	Dataset / sample scale	Imbalance handling	Reported diagnostic performance	Main limitation
Ali et al. [8]	SVM + DNN	Equipment fault data	No explicit balancing	High classification capability	Minority-class imbalance not specifically addressed
Uzel et al. [9]	Optimized ANN-RF	Industrial fault data	No explicit balancing	Improved classification after optimization	Training complexity and imbalance remain concerns
Hairani et al. [10]	Modified SMOTE	Imbalanced classification data	SMOTE-based	Improved minority-class recognition	Synthetic samples may not fully represent difficult boundaries
Wang et al. [14]	Adaptive AdaBoost	Fault classification data	Weight adaptation	Improved classification performance	Does not jointly address sample imbalance and nonlinear representation
Li et al. [16]	Residual-based diagnosis	Power equipment data	No explicit balancing	Effective fault isolation	Relies on residual construction and system-specific information
Karamichailidou et al. [18]	RBF network	Classification data	No explicit balancing	Improved nonlinear approximation	Minority-class recognition and operating-condition robustness not emphasized
SSAR	SVM-SMOTE + AdaBoost-RBF	CWRU-CB: 500 training samples; NASA-Relay: 420 training samples	SVM-guided SMOTE	98.24% diagnostic accuracy; 96.48% F1-score	Further field validation is required

-weight transfer–nonlinear diagnosis," rather than simply concatenating existing algorithms. Compared to existing models—which predominantly involve single-algorithm optimizations or unadapted algorithm stacks—the SSAR model utilizes SVM decision boundaries to guide SMOTE oversampling, preventing the generation of invalid samples, and employs AdaBoost's dynamic weights to drive RBF neural network training, thereby enhancing noise resistance. The dual-module standardized data interface ensures reproducible and transferable results. To clarify the differences between existing methods and the proposed SSAR framework, representative studies are comparatively analyzed in terms of methodology, advantages, limitations, and unresolved research gaps, as summarized in Table 1.As summarized in Table 1, previous studies have generally focused on individual aspects of fault diagnosis, including classifier optimization, nonlinear feature modeling, or class-imbalance mitigation. However, conventional oversampling methods may generate synthetic samples without explicitly accounting for the classifier's decision boundary, while standalone classifiers or fixed ensemble settings may not adequately address samples of different classification difficulty. In addition, reported evaluations are often tied to specific datasets or a limited set of metrics, leaving the combined effects of class imbalance, minority-class recognition, computational efficiency, and operating-condition variation insufficiently characterized. In the quantitative comparison reported in this study, SSAR achieved a diagnostic accuracy of 98.24% and an F1-score of 96.48%, with accuracy exceeding that of the SVM-SMOTE baseline (92.47%) by 5.77 percentage points. These results provide initial evidence supporting the integration of boundary-guided oversampling and adaptive nonlinear classification, although further validation under realistic field conditions remains necessary.

Based on the identified research gap, the main contributions of this study are summarized as follows:

1. A sequential dual-module diagnostic framework is developed. The proposed SSAR model connects SVM-SMOTE-based data preprocessing with AdaBoost-RBF-based fault classification, establishing a unified workflow for low-voltage switch fault diagnosis under class-imbalanced conditions.
2. An integrated imbalance-handling and adaptive classification strategy is implemented. SVM decision-boundary information is used to guide minority-class sample generation, while AdaBoost adaptively

updates sample weights to support the subsequent RBF-based classification process.

3. The diagnostic performance of the framework is evaluated from multiple perspectives. Experiments on the CWRU-CB, NASA-Relay, and self-generated datasets assess data processing, classification performance, computational efficiency, and robustness under specified operating-condition variations. Component-wise comparisons are also conducted to examine the contribution of the integrated modules.

2. PROPOSED SSAR MODEL

This section presents the proposed SSAR framework for low-voltage switch fault diagnosis. The framework consists of two sequential modules: SVM-SMOTE for minority-class sample balancing and AdaBoost-RBF for adaptive nonlinear fault classification. Section 2.1 describes the SVM-SMOTE module, while Section 2.2 presents the AdaBoost-RBF diagnostic module and its integration into the complete SSAR model.

2.1. Design of a low-voltage switch fault data balancing and collection algorithm based on SVMSMOTE

Due to the limited amount of information available for low-voltage switch data collection and the need to further expand the boundary samples of resistance background classification, traditional fault data collection methods have problems such as single dimension, low adoption frequency, and inability to capture special features. The SVM algorithm is capable of analyzing multidimensional complex data and addressing nonlinear challenges. It improves the precision and efficiency of data classification and increases the clarity and robustness of the classification boundary [21–22]. In addition, the SMOTE algorithm has the ability to learn the changes in fault features and to mine deep fault patterns, which effectively balances the problem of class collection and expands the fault information content [23]. Therefore, the SVM and SMOTE algorithms are integrated to develop the SVM-SMOTE algorithm, which aims to enhance the overall generalization performance of the algorithm and improve the fault recall rate. The SVM algorithm makes full use of valid fault data, explores deeper fault feature space data, avoids data overfitting, and improves algorithm robustness [24]. The empirical risk and generalization bound of the SVM are formulated according to statistical learning theory [7], as shown in Eq. (1). The corresponding hyperplane, primal optimization problem, and dual formulation are presented in Eqs. (2)–(4).

$$\begin{cases} remp(\omega) = \frac{1}{n}\sum_{i=1}^{n} L(y_i, f(x_i;\omega)) \\ R(\omega) \le remp(\omega) + \varphi(\frac{n}{h}) \\ \varphi(\frac{n}{h}) = \sqrt{\frac{h\left[ln\frac{2N}{h}+1\right]+ln\left(\frac{4}{n}\right)}{N}} \end{cases} \tag{1}$$

In Equation (1), $remp(\omega)$ represents the empirical risk, $R(\omega)$ represents the structural risk, and $\varphi(\frac{n}{h})$ represents the confidence risk, As the complexity of the hypothesis space increases, the confidence risk generally increases, resulting in a looser generalization bound; therefore, controlling hypothesis-space complexity is important for maintaining generalization performance. N is the number of training samples, n represents the number of training samples, y_i represents the true value, $f(x_i;\omega)$ represents the predicted value, h represents the complexity of the hypothesis space, L represents the loss function, and i represents the sample index. The classification hyperplane process of the SVM algorithm is shown in Equation (2).

$$y = wx_i + b \tag{2}$$

In Equation (2), w is the weight vector, and b represents the threshold vector. The calculation equation for solving the optimal hyperplane of the algorithm is shown in Equation (3).

$$\begin{cases} min\, J = \frac{\|w\|^2}{2} + C\sum_{i=1}^{n}\xi_i \\ s.t.\, y_i(wx_i + b) \ge 1 - \xi_i, \xi_i > 0, i = 1,2,\cdots,n \end{cases} \tag{3}$$

In Equation (3), C represents the penalty factor, and ξ_i represents the slack variable, J represents the objective function, n represents the total number of training samples. The process of introducing the Lagrange multiplier to transform the problem into a dual form is shown in Equation (4).

$$\begin{cases} max\, Q(\alpha) = \sum_{i=1}^{n}\alpha_i - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\alpha_j y_i y_j \langle x_i x_j\rangle \\ s.t. \sum_{i=1}^{n}\alpha_i y_i = 0, \alpha_i > 0, i = 1,2,\cdots,n \end{cases} \tag{4}$$

In Equation (4), α_i and α_j represent the Lagrange multipliers, $Q(\alpha)$ represents the dual objective function. Since SVM alone does not

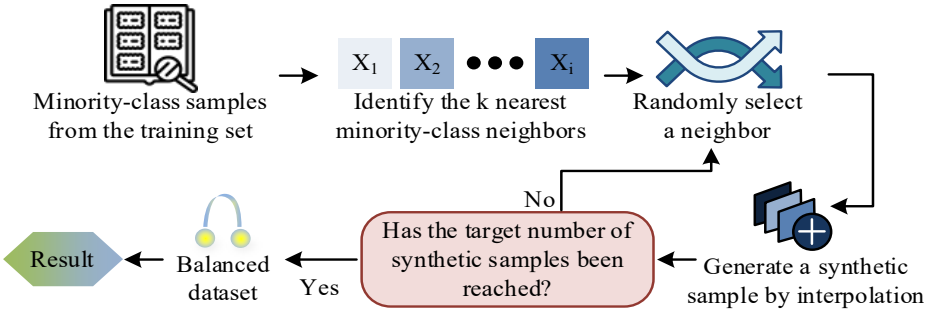


Fig. 1. SMOTE algorithm data processing flow

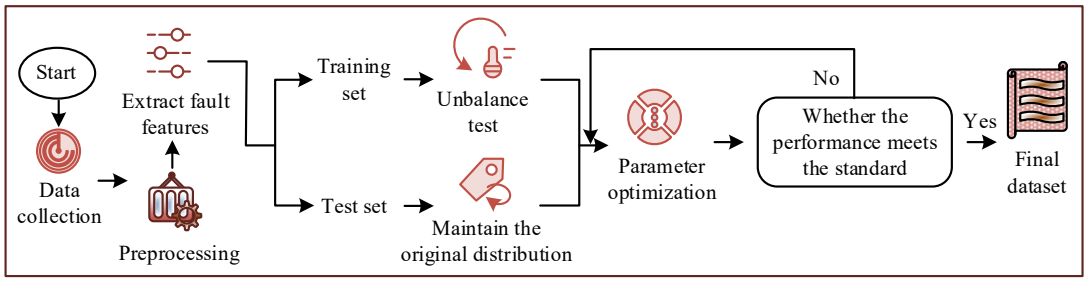


Fig. 2. SVM-SMOTE data collection process

resolve class imbalance, SMOTE is introduced to generate additional minority-class samples. The basic oversampling procedure is illustrated in Fig. 1.

As shown in Fig. 1, SMOTE takes minority-class samples from the training set as input. For each selected sample, the algorithm identifies its k nearest minority-class neighbors using the distance metric defined. One neighbor is then randomly selected, and a synthetic sample is generated by interpolation. The process is repeated until the target number of synthetic samples is reached. The generated samples are then combined with the original training data to form a balanced training set. The test set retains its original class distribution and is excluded from oversampling to prevent data leakage. Figure 1 illustrates the basic SMOTE sample-generation procedure, whereas Fig. 2 presents the SVM-SMOTE data-processing workflow. Following the original SMOTE formulation [11], the distance between neighboring samples and the interpolation-based generation of synthetic samples are defined in Eqs. (5) and (6), respectively.

$$\begin{cases} d_n = \sqrt{\sum_{i=1}^{n}(x_{mn} - x_{nn})^2} \\ d_n = \frac{x_{mn} \cdot x_{nn}}{\|x_{mn}\| \|x_{nn}\|} \end{cases} \quad (5)$$

In Equation (5), x_{mn} represents the sample, x_{nn} represents any neighboring point of the sample, and d_n represents the distance between the sample and any neighboring point. The equation for synthesizing new samples in the algorithm is shown in Equation (6).

$$x_{new} = x_{mm} + \tau * (x_{nn} - x_{mm}) \quad (6)$$

In Equation (6), x_{new} is the new sample, τ is a constant, and $\tau \in (0,1)$. After optimization through the SMOTE, the SVM-SMOTE algorithm has the ability to analyze complex data and resist noise interference. The data collection process of the algorithm is shown in Fig. 2.

Figure 2 illustrates the integrated SVM-SMOTE workflow. The original data are first divided into training and testing sets. The SVM boundary is then estimated from the training data, and SMOTE is applied only to the training set to generate additional minority-class samples. The testing set retains its original distribution to prevent data leakage. The balanced training data are subsequently used for model training and parameter optimization until the stopping criterion is met. The decision function for solving the quadratic convex optimization problem of the SVM-SMOTE algorithm is shown in Equation (7).

$$f(x) = sgn[\sum_{i=1}^{l} \alpha_i^* y_i \langle x^* x_i \rangle + b^*] \quad (7)$$

In Equation (7), sgn represents the sign function used to determine the sample category, and b^* represents the classification threshold. The SVM-SMOTE module and the AdaBoost-RBF module adopt an offline batch serial interaction mode without online iterative feedback. The specific interaction rules are defined as follows:

(1) SVM-SMOTE performs one-time oversampling only on the training set and outputs a fixed training dataset with balanced classes. The test set is not involved in oversampling throughout the process and retains its original class distribution to avoid data leakage.

(2) The balanced training set serves as the only input to the AdaBoost-RBF module for initializing sample weights and training RBF weak classifiers. The two modules are connected through a unified 24-dimensional feature vector interface with consistent feature space.

(3) The training process of the AdaBoost-RBF module does not reversely adjust the sampling strategy of SVM-SMOTE. The two modules are executed serially in the order of preprocessing-classification.

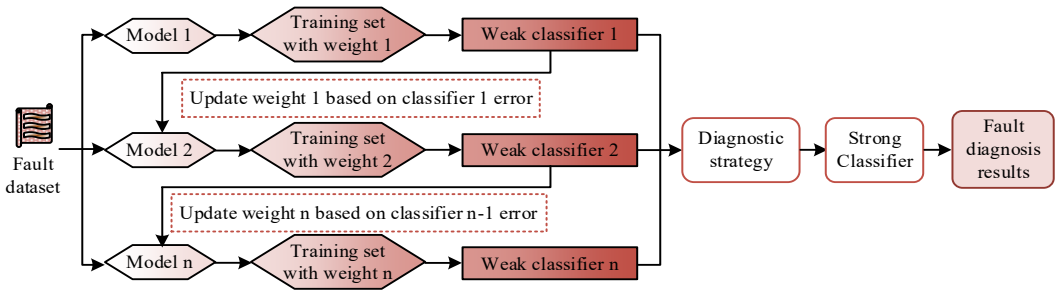


Fig. 3. Fault identification process of AdaBoost algorithm

2.2. Construction of fault diagnosis model introducing AdaBoost-RBF

Following the completion of fault data collection and classification, the SVM-SMOTE algorithm establishes a robust baseline for enhancing diagnostic accuracy and solves the problems of missing fault data and incorrect labels. However, in low-voltage switch fault diagnosis, challenges such as reduced diagnostic accuracy caused by sensor and environmental noise still exist. The AdaBoost algorithm, as a strong classifier, focuses on and assigns different weights to samples, improving the diagnostic accuracy and predictive strength of the model [25]. Meanwhile, the RBF algorithm has excellent ability to process nonlinear data, provides clearer boundary decisions, and enhances the diagnostic efficiency of the model [26]. Recent studies have further demonstrated the effectiveness of hybrid intelligent methods in power equipment fault diagnosis. For example, hybrid approaches combining signal decomposition and machine learning have been applied to low-voltage arc fault identification [27], while comparative studies have explored machine-learning-based fault diagnosis in power systems [28]. These studies demonstrate the potential of hybrid data-driven strategies, but challenges remain in simultaneously addressing imbalanced fault samples and adaptive classification within low-voltage switch diagnosis. Therefore, this paper combines the AdaBoost algorithm with the RBF algorithm to form the AdaBoost-RBF algorithm, which has advantages such as noise resistance, precise fault feature identification, and fast data analysis. The process of fault recognition using the AdaBoost algorithm is shown in Fig. 3.

As shown in Fig. 3, the processed fault feature dataset enters the AdaBoost algorithm for training. The AdaBoost algorithm assigns equal initial weight values to each training sample and inputs them into the corresponding weak classifiers for analysis. During weak classifier analysis, the weight of a correctly identified sample is reduced for the subsequent training iteration. If the sample is misclassified, the algorithm increases its feature weight and inputs it back into the training set. After multiple weak classifiers process the fault feature data, the results are sent to the strong classifier to output the final fault diagnosis classification and complete the training. The AdaBoost sample-weight initialization, weighted error calculation, classifier-weight calculation, and sample-weight update follow the standard boosting formulation [15], as expressed in Eqs. (8)–(11).

$$D(k) = (w_{k1}, w_{k2}, \cdots w_{km}); w_{1i} = \frac{1}{m}; i = 1,2, \cdots m \quad (8)$$

In Equation (8), w_i denotes the weight assigned to the i-th training sample, and m represents the total number of samples in the training dataset. The data are weighted in the k -th weak classifier, but the error rate of misclassified samples is shown in Equation (9).

$$e_k = P(G_k(x_i) \neq y_i) = \sum_{i=1}^{m} w_{ki} I(G_k(x_i) \neq y_i) \quad (9)$$

In Equation (9), $G_k(x_i)$ represents the weak classifier, e_k represents the weighted error rate, P represents the probability of incorrect prediction of

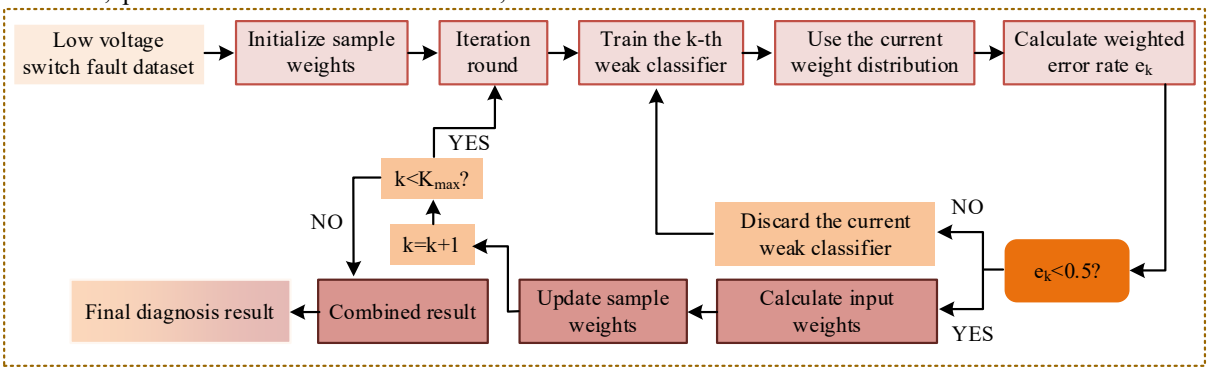


Fig. 4. Flowchart of the AdaBoost-RBF fault diagnosis algorithm

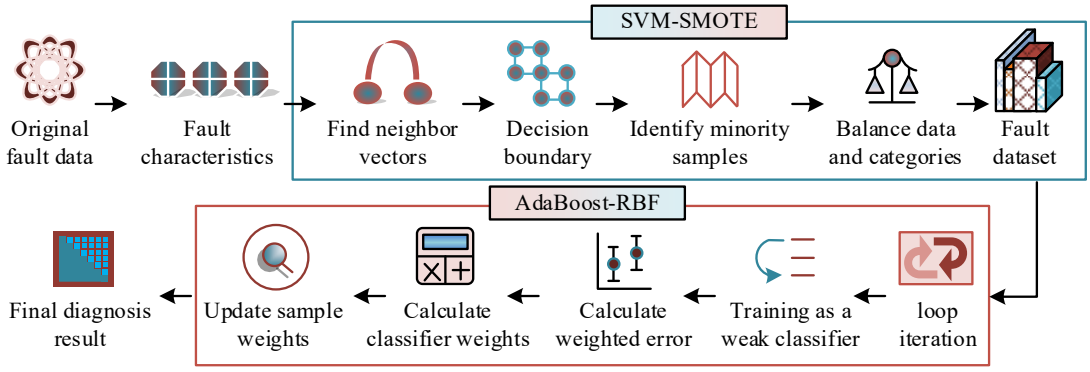


Fig. 5. SSAR model low-voltage switch fault diagnosis process

the weak classifier, y_i represents the i -th true label in the training samples, and I represents the indicator function. The weight coefficient of the fault data on the k -th weak classifier is shown in Equation (10).

$$a_k = \frac{1}{2} log \frac{1-e_k}{e_k} \tag{10}$$

In Equation (10), a_k represents the weight coefficient. The sample weight coefficient of the $k+1$-th weak classifier is shown in Equation (11).

$$w_{k+1,i} = \frac{w_{ki}}{Z_k} e^{-a_k y_i G_k(x_i)} \tag{11}$$

In Equation (11), Z_k represents the normalization factor. Based on the fault diagnosis and classification process of AdaBoost, combining the RBF algorithm's efficient data processing and accurate feature recognition enhances the overall diagnostic accuracy and efficiency of the model for low-voltage switches [29]. The RBF algorithm analyzes fault data mainly through its regional neural network processing. It performs nonlinear transformation of nodes and projects the input spatial data into a higher-dimensional space. It then detects whether the input samples fall within the range of typical fault patterns, performs final feature weighting, and obtains the fault diagnosis result. The Gaussian basis functions and network output are formulated according to the standard RBF network structure [19], as shown in Eqs. (12) and (13).

$$\begin{cases} h_j = exp(\frac{\|x^* - c_j\|^2}{2b_j^2}) \\ f = W^{*T} h(x^*) + \varepsilon \end{cases} \tag{12}$$

In Equation (12), x^* represents the network input, W^{*T} represents the ideal weight of the neural network, j represents the j -th node in the hidden layer, c_j represents the coordinate vector of the j -th neuron center, which reflects the sensitivity of the input features in the studied system, h_j represents the algorithm output, b_j represents the width of the j -th neuron, which reflects the projection ability of the network to the input, and ε represents the approximation error. The output of the neural network is shown in Equation (13).

$$\hat{f}(x^*) = \hat{W}^T h(x^*) \tag{13}$$

In Equation (13), $\hat{f}(x^*)$ represents the neural network output. After combining the nonlinear data fitting ability of the RBF algorithm with the accurate classification ability of the AdaBoost algorithm, the AdaBoost-RBF fault diagnosis algorithm is obtained. The process of diagnosing low-voltage switch faults using this algorithm is shown in Fig. 4.

As shown in Fig. 4, the AdaBoost-RBF algorithm first initializes the fault dataset and assigns an initial weight to each sample. The current weight distribution is then used to train the k-th RBF weak classifier, followed by calculation of its weighted error rate e_k. The condition e_k<0.5 is not guaranteed for every candidate weak classifier but is used as the acceptance criterion for determining whether the current classifier provides better-than-random classification performance. If $e_k \geq 0.5$, the current weak classifier is discarded and does not participate in the subsequent weight update. If $0<e_k<0.5$, the classifier weight is calculated, the sample weights are updated, and the iteration counter is increased. The process continues while $k < K_{max}$; when the preset maximum number of iterations is reached, the retained weak classifiers are combined to obtain the final fault diagnosis result. Using the fault data collected and classified by the SVM-SMOTE algorithm as the basis for diagnosis, and combining the AdaBoost-RBF algorithm to identify fault features and analyze data, this study constructs the final low-voltage switch fault diagnosis model, named SVM-SMOTE-AdaBoost-RBF (SSAR) model. The application process of the SSAR model is shown in Fig. 5.

As illustrated in Fig. 5, raw sensor data from the low-voltage switches undergo preprocessing for fault feature extraction. These features are channeled into the SVM-SMOTE module, where the SVM identifies decision boundaries and isolates samples that exhibit high classification difficulty. Then, the SMOTE algorithm performs oversampling on minority samples to alleviate data imbalance problems and enhance the model's ability to identify minority fault types. The resulting weighted fault feature dataset is subsequently transmitted to the backend, where it serves as an offline fixed dataset for the AdaBoost-RBF module. Through iterative cycles, the training weights of samples on the weak classifier are dynamically adjusted, while sample weight errors are calculated to update the weights in real time. Finally, the weighted strong classifier results are output, including fault types with corresponding confidence levels, completing the low-voltage switch fault diagnosis task.

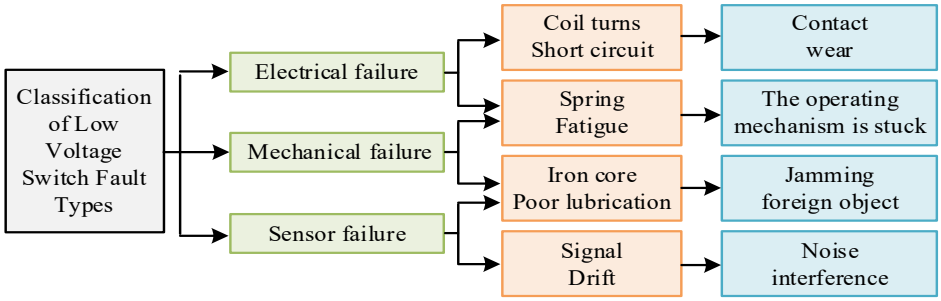


Fig. 6. Classification framework for common faults in low-voltage switchgear

Pseudocode of the proposed SSAR model

Input: Original fault dataset \(D\); number of nearest neighbors \(k\); SVM parameters; AdaBoost-RBF parameters; maximum iteration limit \(K_{\max}\).

Output:Trained SSAR model \(F\); predicted fault labels \(\hat{y}\).

1. Extract the predefined feature vectors from the raw sensor signals.
2. Split the original dataset into training and test sets.
3. Train the SVM on the training set and identify minority-class samples near the decision boundary.
4. Apply SMOTE with the specified \(k\) to the training set only, obtaining the balanced training set \(D_{\mathrm{bal}}\). Keep the test set unchanged.
5. Initialize the sample weights for AdaBoost-RBF.
6. For\(k=1,\ldots,K_{\max}\) do:
   Train the current RBF weak classifier using the weighted training samples.
   Calculate its weighted error rate \(e_k\).
   If the current weak learner satisfies the acceptance criterion, calculate its classifier weight, update and normalize the sample weights, and retain the learner.
   Otherwise, discard the current weak learner without updating the sample weights.
   Stop the boosting process if the predefined termination criterion is met.
7. Combine the retained weak classifiers to construct the final AdaBoost-RBF classifier.
8. Apply the trained SSAR model to the unchanged test set and obtain the predicted fault labels.
9. Evaluate the diagnostic performance using the specified metrics.
10. Return \(F\) and \(\hat{y}\).

To elucidate the fault categories and diagnostic methodology, a comprehensive classification framework is presented for common low-voltage switch faults (electrical faults, mechanical faults, and sensor faults), as shown in Fig. 6.As depicted in Fig. 6, the investigated faults are classified into three primary domains: electrical faults (inter-turn short circuits in coils, contact wear), mechanical faults (spring fatigue, inadequate core lubrication), and sensor faults (signal drift, noise interference). The bolded fault types are included in the study's proprietary dataset. The research also classifies existing fault diagnosis methods based on their technical approaches, as illustrated in Fig. 7.

As shown in Fig. 7, fault diagnosis methods can be categorized into five major types based on their technical approaches: signal processing techniques, data-driven and artificial intelligence methods, model-based approaches, classical diagnostic frameworks, and advanced intelligent methods. The proposed SSAR model integrates SMOTE for data augmentation, SVM for machine learning classifiers, AdaBoost for ensemble learning, and RBF neural networks, representing an interdisciplinary approach that combines data-driven and advanced intelligent methodologies.

3. EXPERIMENTAL EVALUATION

The experimental evaluation is organized into three stages: validation of SVM-SMOTE for data balancing, evaluation of AdaBoost-RBF for fault classification, and comprehensive assessment of SSAR through comparative, robustness, and component-wise analyses.

3.1 Evaluation of SVM-SMOTE

To verify the effectiveness of the SVM-SMOTE algorithm for low-voltage switch fault data collecti on, this study conducted comparative experiments w ith three data collection algorithms: Fast Fourier Tra nsform (FFT), Tomek Links (TLs), and Empirical M ode Decomposition (EMD). The experimental envir onment was set up as follows: the operating system was Windows 10, configured with Intel Core i5 CP U, 16 GB memory, 128 GB solid-state drive memor y, Python version 1.2 was used for editing, and the n umber of training iterations was set to 100. The tap changer had a rated voltage of 1100 kV, r ated

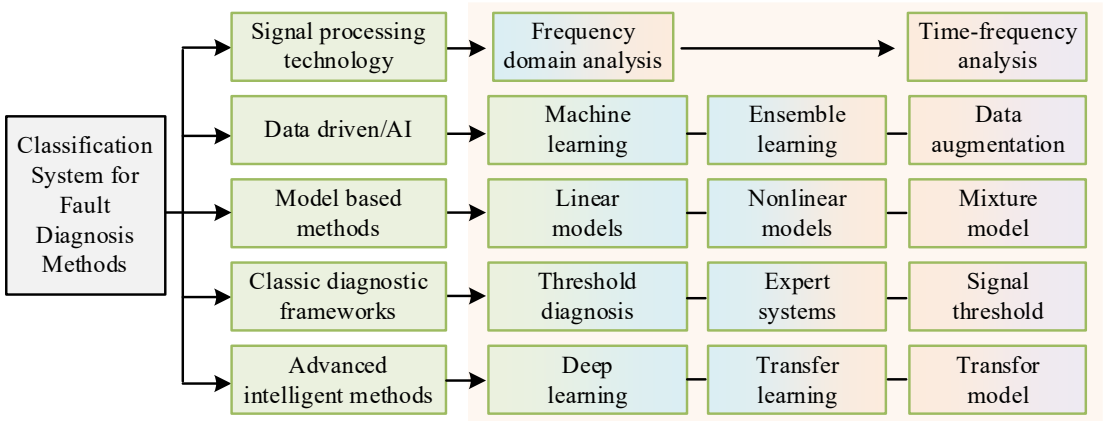


Fig. 7. Existing fault diagnosis methods

current of 6300 A, tapping range of ±5×2.5%, tapping grade of 9 levels, and voltage difference between switches of 0.124 kV. The experimental datasets used were the Case Western Reserve University Circuit Breaker Data Set (CWRU-CB) and the NASA Relay Aging and Life Prediction Data Set (NASA-Relay). The CWRU-CB dataset includes normal operating conditions and four fault modes: spring fatigue, inadequate core lubrication, contact wear, and inter-turn short circuits. The original training set contains 500 samples, including 300 normal samples, 70 spring fatigue samples, 50 inadequate lubrication samples, 60 contact wear samples, and 20 inter-turn short-circuit samples. All comparative models used the same 24-dimensional feature representation and data partitions. Each experiment was independently repeated 30 times, and the results are reported as mean ± standard deviation. The NASA-Relay dataset contains four degradation states during relay aging, with 420 original training samples. Similarly, SMOTE was applied only to the training set after data partitioning, resulting in 210 samples per class, while the testing set remained unchanged. Feature extraction was performed using time-domain statistical features (mean, variance, peak, and kurtosis), frequency-domain features (FFT peak amplitude), and wavelet packet energy features obtained through three-level decomposition across eight frequency bands, resulting in a 24-dimensional feature vector for each sample. The SVM used an RBF kernel with C=10 and γ=0.1, while SMOTE used k=5. The oversampling ratio was adjusted according to the original class size. AdaBoost employed 50 weak classifiers with a learning rate of 0.1, and the RBF network used 20 hidden nodes with a width parameter of 0.5. All algorithms were run for a maximum of 100 iterations, with early stopping after 10 consecutive rounds without improvement. For methodological consistency, the same 24-dimensional feature representation and corresponding data partition were used for all comparative models. SMOTE was applied only to the training set, whereas the test set retained its original distribution. All experiments were independently repeated 30 times, and the reported results were expressed as mean ± standard deviation. These settings were kept unchanged across comparative experiments to ensure reproducibility and a consistent evaluation basis. This study conducted experiments with the SVM-SMOTE algorithm, FFT, TLs, and EMD algorithms on the CWRU-CB dataset, using data integrity as the evaluation metric. The results are shown in Fig. 8.

As shown in Fig. 8, the SVM-SMOTE algorithm achieved a data collection integrity rate of 96.45%. Among the four algorithms, the TLs algorithm had the lowest integrity rate at 88.12%, which was 8.33% lower than the proposed algorithm. The data integrity rates of the FFT algorithm and EMD algorithm were 1.88% and 3.37% lower than the proposed algorithm, respectively. These results indicate that the SVM-SMOTE algorithm demonstrates superior performance in ensuring data integrity in fault data collection, laying a data foundation for in-depth mining of fault features. This was attributed to the collaborative cooperation between SVM and SMOTE algorithms, which enabled multiple classifiers to make more accurate judgments and improved the integrity and quality of information. To verify the effectiveness of the SVM-SMOTE algorithm in data collection, this study conducted experiments with it and three comparative algorithms on the NASA-Relay dataset to test data recognition accuracy. The experimental results are shown in Fig. 9.

As shown in Fig. 9(a), a recognition accuracy of 98.11% was attained by the SVM-SMOTE algorithm in identifying fault types. In Figs. 9(b), (c), and (d), the recognition accuracy of the FFT algorithm was 91.76%, the recognition accuracy of the EMD algorithm was 90.08%, and the TLs algorithm had the worst recognition performance with an accuracy of 88.23%, which was 9.88% lower than the proposed algorithm. These findings suggest that the SVM-SMOTE algorithm demonstrates superior recognition capabilities. This superiority is attributable to the intelligent partitioning function, which facilitates a more concentrated analysis of critical features, effectively suppressed noise influence, and improved the discrimination of core features.

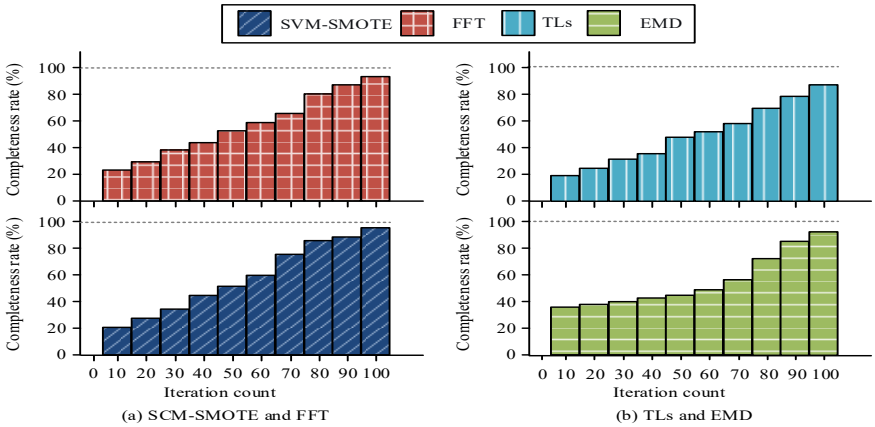


Fig. 8. Comparison of data integrity test results

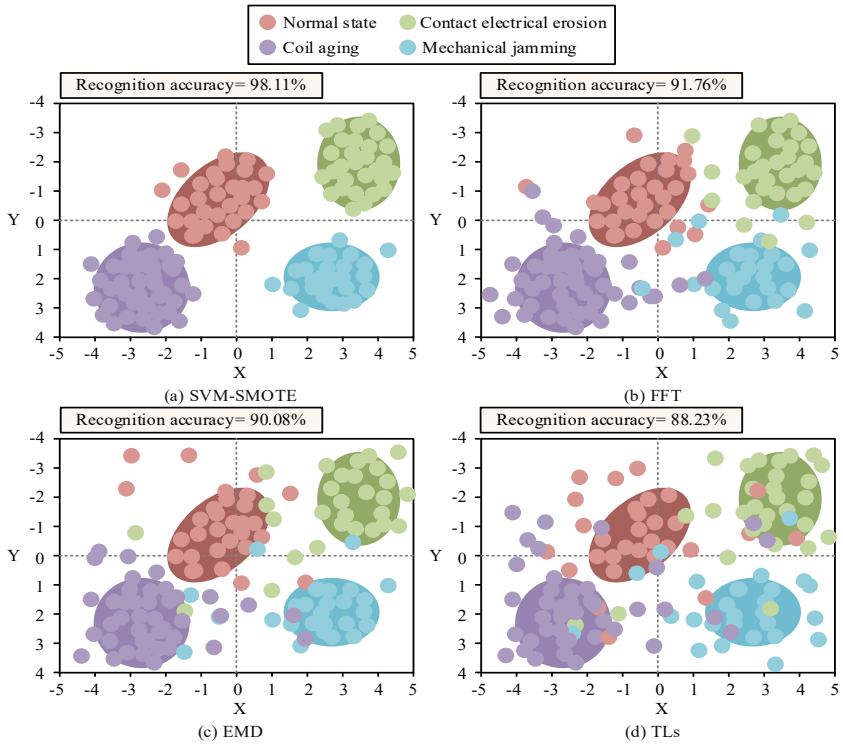


Fig. 9. Comparison of data recognition accuracy results

3.2. Evaluation of AdaBoost-RBF

After verifying the advantages of the SVM-SMOTE algorithm in data collection task integrity and recognition accuracy, this study further analyzed the performance of the AdaBoost-RBF algorithm in fault diagnosis and conducted comparative experiments with three fault diagnosis algorithms: Naive Bayes Classifier (NBC), Deep Belief Network (DBN), and Gaussian Mixture Model (GMM). The experimental environment settings were consistent with the parameters used in the SVM-SMOTE algorithm dataset collection performance verification, and the experimental datasets also used the CWRU-CB dataset and NASA-Relay dataset. This study conducted experiments with the AdaBoost-RBF algorithm, NBC algorithm, DBN algorithm, and GMM algorithm on the CWRU-CB dataset for fault diagnosis classification accuracy, where the quantity of each fault type was set to 100, and the classification capability was expressed in the form of confusion matrix. The experimental results are shown in Fig. 10.

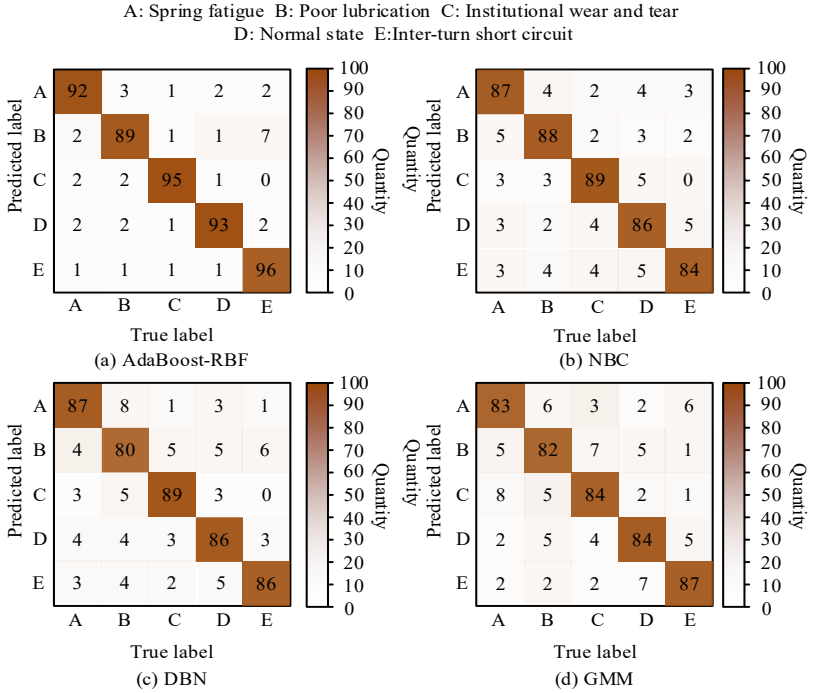


Fig. 10. Comparison of classification accuracy experiment results

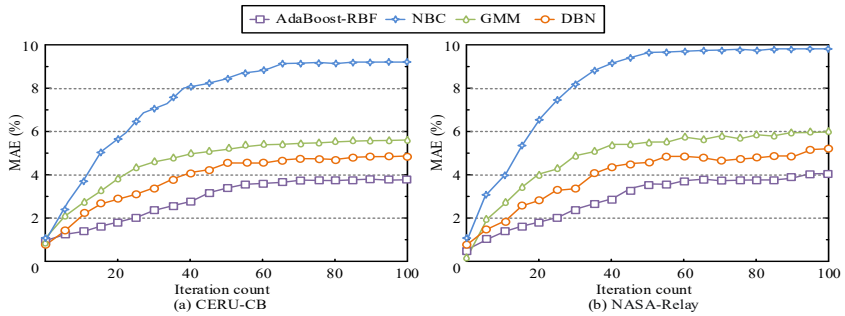


Fig. 11. Comparison of MAE results in terms of accuracy

As shown in Fig. 10(a), the AdaBoost-RBF algorithm achieved an overall classification accuracy of 93.00% across the five fault classes. In Figs. 10(b), (c), and (d), the classification accuracies of NBC, DBN, and GMM were 86.80%, 85.60%, and 84.00%, respectively. Among the comparative models, NBC achieved the highest classification accuracy, but it remained 6.20 percentage points lower than that of AdaBoost-RBF. Notably, the inter-turn short-circuit class (E), which contains relatively few samples, was also effectively identified by AdaBoost-RBF, achieving a recognition rate of 96%. These results indicate that AdaBoost-RBF provides more accurate fault-type identification and superior diagnostic performance than the comparative models. This improvement can be attributed to its adaptive sample-weighting mechanism, which assigns greater weights to difficult-to-classify samples and thereby enhances the overall diagnostic capability. To further verify the diagnostic performance of the proposed algorithm, this study tested it with the above three comparative algorithms on the CWRU-CB dataset and NASA-Relay dataset, using the Mean Absolute Error (MAE) of diagnostic accuracy for analysis. The test results are shown in Fig. 11.

As shown in Fig. 11(a), the MAE of diagnostic accuracy for the proposed algorithm was 3.75%, which was at a relatively low level compared to the other three algorithms. In Fig. 11(b), the diagnostic accuracy MAE of the AdaBoost-RBF algorithm was 4.09%, which was lower than the NBC algorithm's 5.72%, and the proposed algorithm maintained a relatively low error level in tests on both datasets. From the above results, it can be seen that the AdaBoost-RBF algorithm demonstrated higher accuracy and stability when diagnosing different types of faults, with the MAE of both datasets controlled below 5%, possessing precise recognition capability.

3.3. Comprehensive Evaluation of SSAR

After verifying the data collection capability of the SVM-SMOTE algorithm and the diagnostic capability of the AdaBoost-RBF algorithm, this study conducted a comprehensive evaluation of the fault diagnosis performance of the composite model SSAR. This study conducted comparative experiments with the VMD-SVM model composed of Variational Mode Decomposition (VMD) and SVM, the MTF-GAN model composed of Markov Transition Field (MTF) and Graph Attention Network (GAN), and the CNN-BKA model integrating Convolutional Neural Network (CNN) and Bi-directional Kernel Attention (BKA). The experimental environment was consistent with the fault detection settings of the algorithm, and the experimental datasets used were the CWRU-CB dataset, NASA-Relay dataset, and a self-made dataset. The self-made dataset was collected by a fault simulation platform built in the laboratory, containing different levels of load, voltage, and current, with different fault types and sensors set up to simulate real low-voltage switch fault environments. The settings are shown in Table 2.

Table 2. Content of self-made dataset

Dataset	Operational state	Quantity
Training	Normal	300
	Spring fatigue	70
	Poor lubrication of iron core	50
	Contact wear	60
	Short circuit between coil turns	55
Validation	Operational state	30
	Spring fatigue	30
	Poor lubrication of iron core	30
	Contact wear	30
	Short circuit between coil turns	30

As shown in Table 2, the self-built dataset comprises one normal state and four fault states, totaling five operational state categories with 355 labels in total. The training and validation datasets are allocated in a 7:3 ratio. The self-generated dataset utilizes current, voltage, and vibration asd sensors, with a sampling frequency of 10 kHz and a signal length of 512 points. Wavelet packet decomposition and statistical feature extraction were employed, supplemented by SMOTE analysis to compare the sample quantities across categories before and after processing. This study tested the SSAR model with the VMD-SVM model, MTF-GAN model, and CNN-BKA model on the CWRU-CB dataset and self-made dataset, using fault diagnosis time and diagnostic accuracy as evaluation metrics. The test results are shown in Fig. 12.

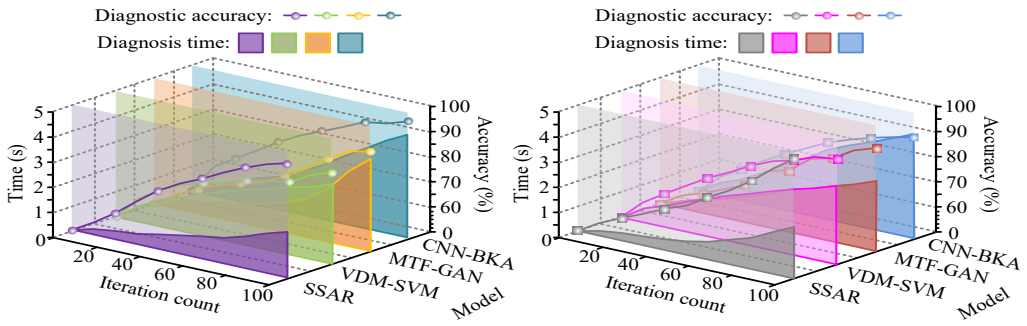


Fig. 12. Comparison of fault diagnosis time and diagnostic accuracy

Table 3. Model indicator test results (Mean ± Std, %, N=30)

Dataset	Model	Recall (%)	F1 (%)	Precision (%)
Self-made	SSAR	95.98 ± 0.32	96.48 ± 0.28	97.11 ± 0.35
	VMD-SVM	90.28 ± 0.87*	91.14 ± 0.76*	90.35 ± 0.92*
	MTF-GAN	88.57 ± 1.02*	87.52 ± 1.15*	86.93 ± 1.08*
	CNN-BKA	82.25 ± 1.24*	82.19 ± 1.31*	83.34 ± 1.19*
CWRU-CB	SSAR	96.46 ± 0.41	95.99 ± 0.37	97.13 ± 0.44
	VMD-SVM	91.02 ± 0.92*	90.05 ± 0.88*	89.68 ± 0.95*
	MTF-GAN	88.74 ± 1.11*	89.69 ± 1.03*	90.01 ± 1.07*
	CNN-BKA	84.56 ± 1.18*	83.57 ± 1.26*	86.45 ± 1.13*
NASA-Relay	SSAR	96.83 ± 0.36	96.85 ± 0.33	95.84 ± 0.39
	VMD-SVM	90.08 ± 0.94*	91.12 ± 0.85*	90.03 ± 0.91*
	MTF-GAN	88.59 ± 1.05*	87.36 ± 1.12*	89.78 ± 1.01*
	CNN-BKA	84.22 ± 1.21*	83.48 ± 1.29*	85.96 ± 1.18*

As illustrated in Fig. 12(a), a fault detection latency of 1.79 s was recorded for the SSAR model, yielding a diagnostic accuracy of 96.13%. The VMD-SVM model exhibited the lowest performance, despite also utilizing the SVM algorithm, its diagnostic latency was 1.49 s higher than that of the proposed model and had an 8.91% decrease in diagnostic accuracy. In Fig. 12(b), the SSAR model required 2.08 s with a diagnostic accuracy of 98.24%. Compared to the CNN-BKA model, the detection time was shortened by 3.16 s and the diagnostic accuracy was improved by 7.18%. These results demonstrate that the SSAR model outperforms existing models in both diagnostic efficiency and accuracy. The model ensured detection precision while controlling processing time, achieving an excellent balance between precision and efficiency. Finally, to validate the enhanced diagnostic performance of the proposed model, it was further tested against the three comparative models on the self-generated dataset, CWRU-CB dataset, and NASA-Relay dataset. Performance was evaluated using three metrics: diagnostic recall rate, F1 score, and precision. Additionally, statistical significance tests were conducted, with all experiments independently repeated 30 times; results are presented as the mean of 30 repetitions. Paired t-tests were performed on key metrics between the SSAR model and each comparative model, using a significance level of $p<0.05$, and results are presented as mean ± standard deviation. The experimental results are shown in Table 3.

Note: * indicates statistically significant differences compared to the SSAR model (paired t-test, $p<0.01$). For a fair comparison, all baseline models and SSAR were evaluated using the same datasets and test partitions, with identical evaluation metrics and 30 independent repetitions.

As shown in Table 3, the SSAR model demonstrated significantly higher recall rate, F1 score, and precision across all three datasets compared to the control model. The recall rate exceeded 95% across all datasets, with the F1 score consistently maintained above 95%, and the highest precision reached 97.13 ± 0.44%. From the above results, it can be seen that the SSAR model maintained highly accurate fault detection capability when facing different experimental environments, controlled the occurrence of errors, and while achieving wide-range coverage inspection, the missed detection rate and false alarm rate were effectively suppressed. It was able to provide highly reliable diagnostic results and was a fault detection model with excellent and robust comprehensive performance. To further assess the robustness of SSAR under practical operating variations, additional experiments were conducted under electrical noise and changes in voltage, current, load, and operating frequency. The nominal operating condition was used as the reference, while controlled perturbations were introduced to each factor individually. The results are summarized in Table 4.As shown in Table 4, SSAR maintained stable diagnostic performance under moderate variations in voltage, current, load, and frequency. The most noticeable degradation occurred under strong electrical noise at SNR = 10 dB, where the F1-score decreased to 87.84%. Nevertheless, the model retained a relatively high

Table 4. SSAR performance under different operating conditions (Mean ± Std, %, N=30)

Operating condition	Setting	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Electrical noise	SNR = 30 dB	91.84 ± 0.71	91.32 ± 0.76	90.95 ± 0.83	91.13 ± 0.78
	SNR = 20 dB	90.76 ± 0.84	90.21 ± 0.89	89.87 ± 0.94	90.04 ± 0.91
	SNR = 10 dB	88.63 ± 1.12	88.05 ± 1.18	87.64 ± 1.24	87.84 ± 1.21
Voltage	−10%	91.52 ± 0.77	91.03 ± 0.82	90.66 ± 0.86	90.84 ± 0.84
	Nominal	92.47 ± 0.68	91.84 ± 0.72	91.23 ± 0.79	91.53 ± 0.75
	+10%	91.36 ± 0.81	90.82 ± 0.87	90.41 ± 0.91	90.61 ± 0.89
Current	−10%	91.28 ± 0.82	90.76 ± 0.86	90.32 ± 0.92	90.54 ± 0.89
	Nominal	92.47 ± 0.68	91.84 ± 0.72	91.23 ± 0.79	91.53 ± 0.75
	+10%	91.43 ± 0.79	90.95 ± 0.83	90.57 ± 0.88	90.76 ± 0.85
Load	80% rated load	90.97 ± 0.88	90.41 ± 0.93	89.98 ± 0.97	90.19 ± 0.95
	100% rated load	92.47 ± 0.68	91.84 ± 0.72	91.23 ± 0.79	91.53 ± 0.75
	120% rated load	91.08 ± 0.91	90.53 ± 0.96	90.06 ± 1.01	90.29 ± 0.98
Frequency	−5%	91.31 ± 0.80	90.84 ± 0.85	90.43 ± 0.89	90.63 ± 0.87
	Nominal	92.47 ± 0.68	91.84 ± 0.72	91.23 ± 0.79	91.53 ± 0.75
	+5%	91.19 ± 0.83	90.71 ± 0.88	90.28 ± 0.94	90.49 ± 0.91

Table 5. Component-wise comparison of the SSAR model

Model	Recall (%)	F1 (%)	Precision (%)
SVM	88.74 ± 0.91	87.92 ± 0.96	89.36 ± 0.88
SVM-SMOTE	92.63 ± 0.72	92.17 ± 0.68	93.21 ± 0.70
AdaBoost-RBF	93.41 ± 0.61	93.18 ± 0.57	94.05 ± 0.59
SVM + AdaBoost-RBF	94.72 ± 0.49	95.01 ± 0.44	95.36 ± 0.46
SSAR	95.98 ± 0.32	96.48 ± 0.28	97.11 ± 0.35

recognition level under this condition. The limited performance fluctuations under voltage, current, load, and frequency variations indicate that the proposed feature representation and adaptive classification mechanism are relatively insensitive to moderate operating-condition changes, supporting the robustness of SSAR for practical fault diagnosis. To further clarify the contribution of each component to the overall performance, a component-wise comparison of the five model configurations was conducted on the self-made dataset, as shown in Table 5.

As shown in Table 5, the progressive incorporation of the proposed components consistently improved the diagnostic performance. The SVM model achieved Recall, F1, and Precision values of 88.74%, 87.92%, and 89.36%, respectively. After introducing SVM-SMOTE, these values increased to 92.63%, 92.17%, and 93.21%, indicating that boundary-guided sample balancing improved the recognition of minority fault classes. The AdaBoost-RBF model further improved the three indicators to 93.41%, 93.18%, and 94.05%, demonstrating the benefit of adaptive sample weighting and nonlinear classification. Combining SVM with AdaBoost-RBF increased the F1 and Precision to 95.01% and 95.36%, respectively. The complete SSAR model achieved the highest Recall, F1, and Precision of 95.98%, 96.48%, and 97.11%, respectively, confirming that the joint use of SVM-SMOTE and AdaBoost-RBF provides additional performance gains rather than merely increasing model complexity.

Two main challenges are encountered during the implementation of this study. First, rare fault samples of low-voltage switches are scarce, and sample features near the decision boundary are fuzzy. Conventional SMOTE methods tend to generate outlier synthetic samples deviating from the effective distribution, which further degrades classification performance. To address this issue, an SVM decision-boundary-constrained interpolation strategy is adopted to limit synthetic samples within the effective feature space near the classification boundary, improving oversampling quality. Second, the multi-algorithm coupled model has high parameter tuning dimensions and long computational time. A 5-fold cross-validation combined with grid search strategy is adopted, and an early stopping mechanism is introduced to terminate iterations when the validation set accuracy shows no improvement for 10 consecutive rounds. All algorithms are implemented based on Python, with core tool libraries including Scikit-learn, NumPy and SciPy for model construction, numerical calculation and signal processing.

4. CONCLUSION

As a core control component in distribution systems, low-voltage switches face challenges including limited fault samples, weak recognition of

minority classes, and insufficient diagnostic accuracy. To address these issues, a hybrid SSAR model integrating SVM-SMOTE and AdaBoost-RBF was developed for low-voltage switch fault diagnosis. SVM-SMOTE was employed to alleviate sample imbalance and enhance minority-class representation, while AdaBoost-RBF was subsequently used to improve fault-type identification through adaptive sample weighting and nonlinear classification. The proposed framework therefore addresses the combined challenges of imbalanced fault samples, noise interference, and multi-algorithm integration. Experimental results show that SVM-SMOTE achieved a data integrity rate of 96.45% and a fault identification accuracy of 98.11%. The fault diagnosis algorithm achieved a classification accuracy of 93.00%, with an MAE of 3.75%. The SSAR model required 2.08 s for 100 iterations and achieved a diagnostic accuracy of 98.24%, with precision and recall above 95% and an F1-score of 96.48%. These results demonstrate that the proposed framework can effectively improve fault data processing and diagnostic performance, particularly under imbalanced fault-sample conditions. Overall, SSAR shows promising potential for low-voltage switchgear fault diagnosis and provides an effective framework for addressing sample imbalance and insufficient minority-class recognition.

The proposed method has the potential to be generalized to other power equipment and fault diagnosis scenarios. From the methodological perspective, the core logic of SVM-boundary-guided SMOTE oversampling does not depend on the physical characteristics of specific equipment, and is applicable to various fault diagnosis tasks with class imbalance problems, such as motor bearing faults and transformer insulation faults. The adaptive weighting and nonlinear classification mechanism of AdaBoost-RBF is universal for high-dimensional features, and migration can be completed only by adjusting the input feature dimension and base classifier parameters. From the verification perspective, stable and excellent diagnostic performance is achieved on three types of equipment with different structures and fault modes: circuit breakers, relays, and self-developed low-voltage switches, which preliminarily verifies the cross-scenario applicability of the method. Further verification on more power equipment such as transformers and transmission lines will be conducted in future work to expand the application scope of the method.

However, the current validation is primarily based on the available experimental datasets, whose sample sizes and operating-condition coverage remain limited. Future work will expand data collection across additional low-voltage switchgear units, fault categories, and operating conditions, and evaluate SSAR using larger independent datasets and cross-device validation. Further improvements in model robustness and adaptability, together with the development of lightweight architectures for edge-computing platforms, will also be investigated.

Source of funding: *This research received no external funding.*

Author contributions: *Research concept and design, K.L.; Collection and/or assembly of data, F.S.; Data analysis and interpretation, Y.L.; Writing the article, S.X.; Critical revision of the article, S.X.; Final approval of the article, K.L.*

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

REFERENCES

1. Pandey AK, Kishor N, Mohanty SR, Samuel P. Intelligent fault detection and classification for an unbalanced network with inverter-based DG units. IEEE Transactions on Industrial Informatics. 2024;20(5): 7325-7334. https://doi.org/10.1109/TII.2024.3359450.
2. Guerrero JM, Aranzabal I, Gomez-Cornejo J, Eguia P, Zarate S, Platero CA. Ground fault detector for DC microgrids using sequentially-switched grounding connections. IEEE Transactions on Industry Applications. 2024; 60(6): 8888-8900. https://doi.org/10.1109/TIA.2024.3446749.
3. Luo R, Gao H, Peng F, Guo Y. An adaptive low-voltage ride-through method for virtual synchronous generators during asymmetrical grid faults. IEEE Transactions on Sustainable Energy. 2024;15(3):1589-1600. https://doi.org/10.1109/TSTE.2024.3358394.
4. Aluri VRK, Samantaray SR. An improved probe power unit (PPU)-based fault location estimation in low-voltage DC microgrids. IEEE Transactions on Smart Grid. 2024;15(6):5514-5525. https://doi.org/10.1109/TSG.2024.3412591.
5. Kumar VTRP, Arulselvi M, Sastry KBS. Comparative assessment of colon cancer classification using diverse deep learning approaches. Journal of Data Science and Intelligent Systems. 2023;1(2):128-135. https://doi.org/10.47852/bonviewjdsis32021193.
6. Que Z, Lin CJ. One-class SVM probabilistic outputs. IEEE transactions on neural networks and learning systems. 2024;36(4):6244-6256. https://doi.org/10.1109/TNNLS.2024.3395148.
7. Cortes C, Vapnik V. Support-vector networks. Machine Learning. 1995;20(3):273–297. https://doi.org/ 10.1007/BF00994018.
8. Ali L, Javeed A, Noor A, Rauf HZ, Kadry S, Gandomi AH. Parkinson's disease detection based on features refinement through L1 regularized SVM and deep neural network. Scientific Reports 2024; 14(1): 1333-1346. https://doi.org/10.1038/s41598-024-51600-y
9. Uzel H, Özüpak Y, Alpsalaz F, Aslan E. Optimized ANN–RF hybrid model with optuna for fault detection and classification in power transmission systems. Scientific Reports. 2025;16(1). https://doi.org/10.1038/s41598-025-31008-y.
10. Hairani H, Widiyaningtyas T, Prasetya DD. Addressing class imbalance of health data: a systematic literature review on modified synthetic minority oversampling technique (SMOTE) strategies. JOIV: International Journal on Informatics Visualization. 2024;8(3):1310-1318. https://doi.org/10.62527/joiv.8.3.2283.
11. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: Synthetic Minority Over-sampling Technique. Journal of Artificial Intelligence Research. 2002;16: 321–357. https://doi.org/10.1613/jair.953.

12. Nugroho K, Muslikh AR, Irianandа SW, Ojugo AA. Integrating SMOTE-tomek and fusion learning with XGBoost meta-learner for robust diabetes recognition. Journal of Future Artificial Intelligence and Technologies. 2024;1(1):23-38. https://doi.org/10.62411/faith.2024-11.
13. Guemmour MB, Negadi K, Araria R, Bey M. Analysis and fault diagnosis in point absorber wave energy conversion systems using fault tree and Bayesian networks. Diagnostyka. 2025;26(2):2025209. https://doi.org/10.29354/diag/204573.
14. Wang S, Liu W, Yang S, Huang H. An optimized AdaBoost algorithm with atherosclerosis diagnostic applications: adaptive weight-adjustable boosting. The Journal of Supercomputing. 2024;80(9):13187-13216. https://doi.org/10.1007/s11227-024-05951-y.
15. Freund Y, Schapire RE. A decision-theoretic generalization of on-line learning and an application to boosting. Journal of Computer and System Sciences. 1997;55(1):119–139. https://doi.org/10.1006/jcss.1997.1504.
16. Li T, Xia Q, Ouyang Y, Zeng R, Liu Q, Li T. Prospectivity and uncertainty analysis of tungsten polymetallogenic mineral resources in the nanling metallogenic belt, south china: a comparative study of AdaBoost, GBDT, and XgBoost algorithms. Natural Resources Research. 2024;33(3):1049-1071. https://doi.org/10.1007/s11053-024-10321-9.
17. Li M, Yu D, Chen Z, Xiahou K, Ji T, Wu QH. A data-driven residual-based method for fault diagnosis and isolation in wind turbines. IEEE Transactions on Sustainable Energy. 2018;10(2):895-904. https://doi.org/10.1109/TSTE.2018.2853990.
18. Karamichailidou D, Gerolymatos G, Patrinos P, Sarimveis H, Alexandridis A. Radial basis function neural network training using variable projection and fuzzy means. Neural Computing and Applications. 2024; 36(33):21137-21151. https://doi.org/10.1007/s00521-024-10274-3.
19. Moody J, Darken CJ. Fast learning in networks of locally-tuned processing units. Neural Computation. 1989;1(2):281–294. https://doi.org/10.1162/neco.1989.1.2.281.
20. Zhang Y, Cao J, Zhang B, Zheng X, Chen W. A comparative study of different radial basis function interpolation algorithms in the reconstruction and path planning of γ radiation fields. Nuclear Engineering and Technology. 2024;56(7):2806-2820. https://doi.org/10.1016/j.net.2024.02.043.
21. Deb A, Jain AK. An effective data-driven machine learning hybrid approach for fault detection and classification in a standalone low-voltage DC microgrid. Electrical Engineering. 2024;106(5):6199-6212. https://doi.org/10.1007/s00202-024-02334-7.
22. Wahyuni SD, Kusumodestoni RH. Optimalisasi algoritma support vector machine (SVM) dalam klasifikasi kejadian data stunting. Bulletin of Information Technology (BIT). 2024;5(2):56-64. https://doi.org/10.47065/bit.v5i2.1247.
23. Zhou Z, Xu C, Qiao Y, Xiong J, Yu J. Enhancing equipment health prediction with enhanced SMOTE-KNN. Journal of Industrial Engineering and Applied Science, 2024;2(2):13-20. https://doi.org/10.5281/zenodo.10775532.
24. Modhugu VR, Ponnusamy S. Comparative analysis of machine learning algorithms for liver disease prediction: SVM, logistic regression, and decision tree. Asian Journal of Research in Computer Science. 2024;17(6): 188-201. https://doi.org/10.9734/ajrcos/2024/v17i6467.
25. Jasim AA, Hazim LR, Mohammedqasim H, Mohammedqasim R, Ata O, Salman OH. e-Diagnostic system for diabetes disease prediction on an IoMT environment-based hyper AdaBoost machine learning model. The Journal of Supercomputing. 2024;80(11): 15664-15689. https://doi.org/10.1007/s11227-024-06082-0.
26. Lin H, Dai H, Mao Y, Wang L. An optimized radial basis function neural network with modulation-window activation function. Soft Computing. 2024;28(5):4631-4648. https://doi.org/10.1007/s00500-023-09207-4.
27. Li B, Wu J. Low-voltage arc fault identification using a hybrid method based on improved salp swarm algorithm – variational mode decomposition – random forest. IEEE Access. 2024;12:15410-15418. https://doi.org/10.1109/ACCESS.2024.3354177.
28. Venkatachalam Y, Subbaiyan T. Intelligent fault diagnosis in power systems: A comparative analysis of machine learning-based algorithms. Expert Systems with Applications. 2025;265:125945. https://doi.org/10.1016/j.eswa.2024.125945.
29. Al-Haddad LA, Shijer SS, Jaber AA, Ani ST, Zubaidi AA, Abd ET. Application of AdaBoost for stator fault diagnosis in three-phase permanent magnet synchronous motors based on vibration–current data fusion analysis. Electrical Engineering. 2024;106(4):4527-4542. https://doi.org/10.1007/s00202-023-02195-6.
30. Ziad H, Al-Dujaili A, Humaidi AJ. Techniques used in faults detection of permanent magnet synchronous motor: A review. AIP Conference Proceedings. 2024; 3232(1):030004. https://doi.org/10.1063/5.0236292.
31. Faraj HZ, Al-Dujaili AQ, Humaidi AJ. The classification method of electrical faults in permanent magnet synchronous motor based on deep learning. In: 2023 IEEE 11th Conference on Systems, Process & Control (ICSPC). Malacca, Malaysia. 2023. p. 326–331. https://doi.org/10.1109/ICSPC59664.2023.10420154.
32. Ziad H, Al-Dujaili AQ, Humaidi AJ. A comparative study of deep learning efficiency in the classification of electrical faults of permanent magnet synchronous motor. International Review of Applied Sciences and Engineering. 2025;16(2):292–302. https://doi.org/10.1556/1848.2024.00885.



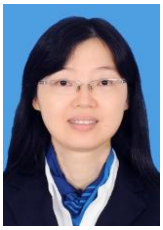
Kaimin LIAN, he was born in Jieyang City, Guangdong Province, People's Republic of China, in November 1986. He holds a bachelor's degree from Zhejiang University. Currently, he works at the Guangdong Power Grid Co., Ltd. Jieyang Jiedong Power Supply Bureau. His research interests include power engineering technology.
e-mail: LXK946824@163.com



Sheng XIE, born in Jieyang City, Guangdong Province, People's Republic of China, in May 1991, he holds a bachelor's degree in Marine Science from Guangdong Ocean University. He currently works at the Guangdong Power Grid Co., Ltd. Jieyang Jiedong Power Supply Bureau. His research interests include power engineering technology.
e-mail: XS45698752@163.com



Feng SUN, he was born in Jieyang City, Guangdong Province, People's Republic of China, in August 1978. He holds a bachelor's degree from Sichuan University. Currently, he works at the Guangdong Power Grid Co., Ltd. Jieyang Jiedong Power Supply Bureau. His research interests include power engineering technology.
e-mail: Sfeng6578962@163.com



Yuanping LUO, she was born in Jieyang City, Guangdong Province, People's Republic of China, in November 1976. She holds a bachelor's degree from South China University of Technology. Currently, she works at the Guangdong Power Grid Co., Ltd. Jieyang Jiedong Power Supply Bureau. Her research interests include infrastructure project management.
e-mail: LWP54863@163.com