

Article

# A Method for Identifying External Short-Circuit Faults in Power Transformers Based on Support Vector Machines

Hao Du <sup>1</sup>, Linglong Cai <sup>2</sup>, Zhiqin Ma <sup>2</sup>, Zhangquan Rao <sup>2</sup>, Xiang Shu <sup>2</sup>, Shuo Jiang <sup>2</sup>, Zhongxiang Li <sup>2</sup>  
and Xianqiang Li <sup>1,\*</sup>

<sup>1</sup> School of Automation, Wuhan University of Technology, Wuhan 430070, China; 334874@whut.edu.cn

<sup>2</sup> Electric Power Research Institute of Guangdong Power Grid Co., Ltd., Guangzhou 510080, China; cailinglong1107@163.com (L.C.); mzhqcqu@163.com (Z.M.); zq\_rao@163.com (Z.R.); shux1173@163.com (X.S.); 13922467757@163.com (S.J.); lizhongxiang\_tbea@126.com (Z.L.)

\* Correspondence: lxq@whut.edu.cn

**Abstract:** Being a vital component of electrical power systems, transformers significantly influence the system stability and reliability of power supplies. Damage to transformers may lead to significant economic losses. The efficient identification of transformer faults holds paramount importance for the stability and security of power grids. The existing methods for identifying transformer faults include oil chromatography analysis, temperature assessment, frequency response analysis, vibration characteristic examination, and leakage magnetic field analysis. These methods suffer from limitations such as limited sensitivity, complexity in operation, and a high demand for specialized skills. In this paper, we propose a method to identify external short-circuit faults of power transformers based on fault recording data on short-circuit currents. It involves analyzing the current signals of various windings during faults, extracting appropriate features, and utilizing a classification algorithm based on a support vector machine (SVM) to determine fault types and locations. The influence of different kernel functions on the classification accuracy of SVM is discussed. The results indicate that this method can proficiently identify the type and location of external short-circuit faults in transformers, achieving an accuracy rate of 98.3%.

**Keywords:** power transformer; external short-circuit fault; fault identification method; SVM



**Citation:** Du, H.; Cai, L.; Ma, Z.; Rao, Z.; Shu, X.; Jiang, S.; Li, Z.; Li, X. A Method for Identifying External Short-Circuit Faults in Power Transformers Based on Support Vector Machines. *Electronics* **2024**, *13*, 1716. <https://doi.org/10.3390/electronics13091716>

Academic Editor: Lazhar Ben-Brahim

Received: 2 April 2024

Revised: 23 April 2024

Accepted: 25 April 2024

Published: 29 April 2024



**Copyright:** © 2024 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

As a pivotal core component of power systems, transformers play a crucial role in the transmission of electrical energy. Transformer malfunctions may lead to the risk of power outages and safety hazards, resulting in economic losses. Therefore, the proper implementation of transformer fault identification is highly important in ensuring the stable operation of power systems, mitigating losses, and optimizing equipment maintenance [1].

According to the different state quantities, the existing fault identification methods for transformers mainly focus on the analysis of dissolved gases in oil, temperature, vibration characteristics, leakage magnetic field, and so on [2–5]. An imprecise probability-based approach for transformer fault identification has been proposed in the literature [6], utilizing the imprecise Dirichlet model and naive credal classifier. In [7], a novel fault identification method was introduced, leveraging a hypersphere multi-class support vector machine (HMSVM) and Dempster–Shafer (D–S) evidence theory (DET). Additionally, Ref. [8] systematically explored and compared interpretation methods for dissolved gas analysis (DGA), encompassing both conventional and intelligent approaches. The authors of [9] introduced a novel multi-input, multi-output polynomial neural network (PNN). Furthermore, in [10], a new method for transformer fault identification was proposed, based on correlation coefficient density clustering.

Huang [11] proposed that during transformer faults, there are unique heat distributions and characteristics. Building on this foundation, a new fault recognition solution

based on the fault overlay method was suggested. The fault recognition capability was achieved by training a convolutional neural network (CNN) model, leading to the successful recognition of faults. In [12], machine learning was utilized to estimate hotspot temperatures, facilitating thermal state monitoring. Transformer fault recognition was ultimately achieved through a thermal anomaly detection algorithm. Cheng [13] established a static aging failure model based on the winding hot spot temperature (HST), utilizing the Weibull distribution and Arrhenius reaction law. The objective of this model was to assess the failure rates and life expectancy of transformers by characterizing the aging mechanism of transformers and estimating the winding HST. The authors of [14] introduced a fault identification method based on the inversion of transformer top oil temperature rise. Transformer fault detection was accomplished in [15] by combining advanced thermal modeling and thermography with image processing methods. This involved comparing top oil and radiator temperatures. Furthermore, in [16], the internal temperature variations of transformers were investigated, leading to the proposal of a sensor-based online decision prediction fault identification method.

A new method for transformer fault identification was proposed in [17]. This method entailed establishing a mathematical model to address the transformer vibration phenomenon and presenting a unique approach for analyzing the vibration spectrum. The authors of [18] presented a unique approach to transformer fault identification for analyzing the vibration characteristics of windings under various fault conditions. The vibration characteristics of transformers under the influence of load current were investigated in [19], leading to the proposal of a novel approach for transformer fault identification. Additionally, the authors of [20] suggested collecting the vibration signal of transformers under fluctuating operating conditions as the characteristic quantity. These features were then processed using the linear discriminant analysis (LDA) algorithm to identify faults.

The authors of [21] proposed the utilization of the magnetic field leakage waveform at the measuring point of the simulated transformer winding, coupled with a convolutional neural network classification model, for identifying transformer faults. Distinctive changes in magnetic flux leakage during fault conditions were identified as representative features through the modeling and analysis of transformers in [22]. In [23], the online analysis of transformer magnetic flux leakage served as the basis for identifying transformer faults. For transformer fault location identification, the authors of [24] detected an asymmetric structure in the distribution of magnetic flux leakage.

The analyses above indicate that the existing methods for transformer fault identification primarily focus on gas analysis, temperature assessment, vibration analysis, magnetic flux leakage, and similar factors. Research on fault identification based on current characteristics during short-circuit processes is relatively limited. In this paper, we propose a method for identifying external short-circuit faults in power transformers based on the fault recording data. This method analyzes the short-circuit current signals of each phase, extracts reasonable feature quantities, and employs the support vector machine (SVM) classification algorithm to determine fault types. Due to the limited availability of actual fault recording data, the classification algorithm may lack the required sample size. In this study, a “field-circuit” coupling model is developed according to a specific power transformer. The short-circuit currents for various external short-circuit fault conditions are computed, serving as data samples. The proposed method’s effectiveness is validated using test samples and fault recording data. This research provides a feasible solution for the precise identification of external short-circuit faults using electrical quantities. It contributes to enhancing the safety and stability of power systems and offers valuable theoretical support for practical applications in electrical engineering.

The rest of this paper is structured as follows. In Section 2, the SVM-based fault identification method is introduced. In Section 3, the data collection procedure is presented, including the establishment and verification of the “field-circuit” coupling model and the computation of short-circuit currents. In Section 4, the algorithmic process of the fault identification method is outlined and the impact of different kernel functions on

identification accuracy is discussed. The results of fault identification are described in Section 5. In Section 6, the paper is summarized.

## 2. Fault Identification Method Based on SVM

Artificial intelligence algorithms have become a prominent research focus in fault identification. Numerous algorithms such as decision trees, Bayes classifiers, neural networks, and SVM have been widely employed in fault identification. SVM is rooted in statistical learning theory, operating on the principle of structural risk minimization. This principle facilitates the derivation of decision rules based on limited sample training, ensuring minimal errors even when applied to independent test sets. Due to its outstanding performance in high-dimensional space and its robust fault classification ability, this paper adopts a multi-class SVM algorithm for external short-circuit fault identification in transformers.

### 2.1. Basic Principle of SVM

A multi-class SVM involves a dataset comprising  $N$  samples, represented as  $X = \{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N), x_n \in R^n, y_n \in (1, 2, \dots, M)\}$ , where  $y_n$  represents the class label for each sample and  $M$  represents the number of classes in the sample data. The goal of the multi-class SVM is to find hyperplanes defined by  $\omega^T x + b = 0$ , where  $\omega$  denotes the normal vector to the hyperplane and  $b$  signifies the distance from the origin to the hyperplane. These hyperplanes should correctly separate the  $N$  samples into  $M$  classes. This scenario is referred to as linear separability. However, in practice, the obtained samples are often not linearly separable. Therefore, a kernel function,  $K(x_i, x_j)$ , is introduced to map these  $N$  samples into a higher-dimensional space. Each sample,  $x_i$ , in this higher-dimensional space, is represented as a feature vector,  $\varphi(x_i)$ . The decision function,  $f(x)$ , is then applied to correctly classify these samples.

The decision function can be expressed as follows:

$$f(x) = \omega^T \varphi(x) + b \tag{1}$$

The kernel function is defined as follows:

$$K(x_i, x_j) = \varphi(x_i)^T \varphi(x_j) \tag{2}$$

For training samples that do not satisfy the constraint, slack variables,  $\xi_i$ , and a penalty factor,  $c$ , are introduced to measure the constraint violation degree for individual samples in the training set. The optimal objective function and constraints are as follows:

$$\begin{cases} \min_{\omega, b, \xi} \left( \frac{1}{2} \omega^T \omega + c \sum_{i=1}^N \xi_i \right) \\ \text{s.t. } y_i (\omega^T \varphi(x_i) + b) \geq 1 - \xi_i, \xi_i \geq 0 \quad i = 1, 2, \dots, N \end{cases} \tag{3}$$

By introducing Lagrange multipliers,  $\alpha_i$ , the dual form of the SVM can be obtained as follows:

$$\begin{cases} \max_{\alpha} \left( -\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j K(x_i, x_j) + \sum_{i=1}^N \alpha_i \right) \\ \text{s.t. } 0 \leq \alpha_i \leq c \quad i = 1, 2, \dots, N \\ \sum_{i=1}^N \alpha_i y_i = 0 \end{cases} \tag{4}$$

The corresponding optimal classification function can be obtained as follows:

$$f(x) = \text{sgn} \left( \sum_{i=1}^N \alpha_i^* y_i K(x_i, x_j) + b^* \right) \tag{5}$$

Commonly employed kernel functions encompass the linear kernel (LN), polynomial kernel (PL), radial basis function kernel (RBF), and sigmoid kernel (SIG). The definitions of each kernel function are as follows:

$$K(x_i, x_j) = x_i^T x_j \quad (6)$$

$$K(x_i, x_j) = (\gamma x_i^T x_j + r)^p, \gamma > 0 \quad (7)$$

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2), \gamma > 0 \quad (8)$$

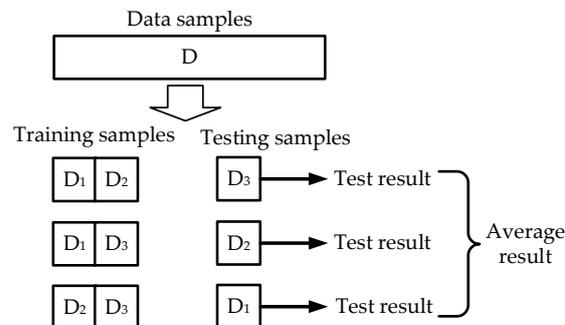
$$K(x_i, x_j) = \tanh(v x_i^T x_j + c) \quad (9)$$

## 2.2. Parameter Optimization Method

The key to using SVM for transformer fault identification is to obtain the corresponding parameters in the decision function,  $f(x)$ . The model exhibits excellent accuracy on both the training and test sets when the parameters are within a reasonable range.

If the model demonstrates high accuracy on the training samples, but low accuracy on the testing samples, it is overfitting. When the model's accuracy is low on both the training and testing samples, it is underfitting.

To improve the accuracy of fault identification, this paper employs K-fold cross-validation (K-CV) to evaluate the performance indicators of the classification model. This approach helps to find the optimal parameter combination. K-CV involves dividing the original data sample into K equal parts. During each iteration, K-1 parts are used for training and the remaining part is utilized for testing. After each validation, an accuracy score representing the classifier's performance is obtained. This process is repeated K times, ensuring that each part of the dataset serves as the validation set once. The overall performance of the classifier under a specific set of parameters in K-CV is then represented by the average of the K accuracy scores. When the highest accuracy score is obtained, its corresponding parameters are considered the optimal parameter combination for the classification model. This study employs 3-fold cross-validation (Figure 1).



**Figure 1.** Schematic diagram of 3-fold cross-validation.

## 3. Data Collection

Existing fault identification methods overly rely on on-site operation and experimental data. When on-site data are insufficient, fault identification is constrained and may struggle to fulfill its purpose. This limitation also affects its effectiveness in giving early warnings of transformer faults. Therefore, we construct a finite element simulation model, using an actual transformer, and verify its correctness. The model is used to simulate various external transformer fault conditions, and the short-circuit currents are obtained as the original data. This approach addresses the challenge of limited on-site data.

### 3.1. Establishment of Transformer Model

In this study, we use three actual power transformers as simulation objects. A field-circuit coupling model of the transformer is established to calculate the short-circuit current.

Figures 2 and 3 illustrate the structural model and the external circuit of the transformer, respectively. The neutral points of the No. 1 and No. 2 units remain ungrounded, while the No. 3 transformer’s neutral points are grounded at both high-voltage and medium-voltage sides. The transformer windings include medium-voltage (MV) winding 1, voltage-regulating (VR) winding, high-voltage (HV) winding, medium-voltage winding 2, and low-voltage (LV) winding components. The high-, medium-, and low-voltage ratings are 220 kV, 115 kV, and 10.5 kV, respectively. Table 1 provides an overview of the electrical parameters of the transformer, including the rated capacity, voltage, current, connection group number, and the number of turns for each winding. Table 2 outlines the structural dimensions of the transformer, encompassing the diameter and window height of the core, center distance, yoke height, winding height, and the diameter range for each winding.

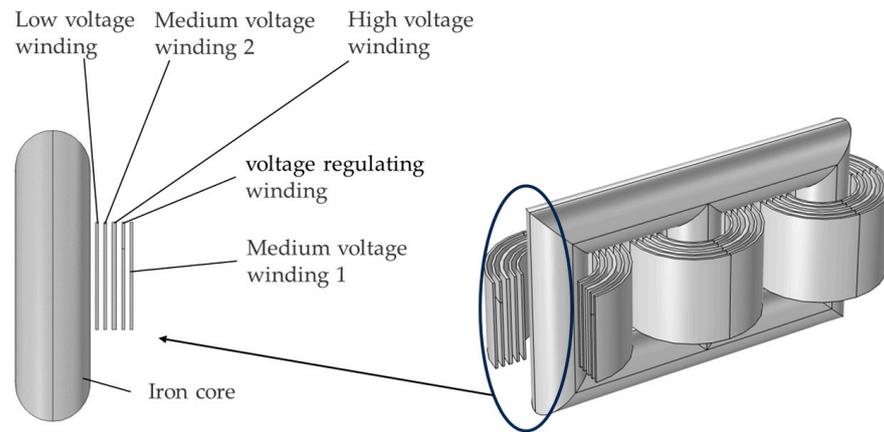


Figure 2. Structural model of transformer.

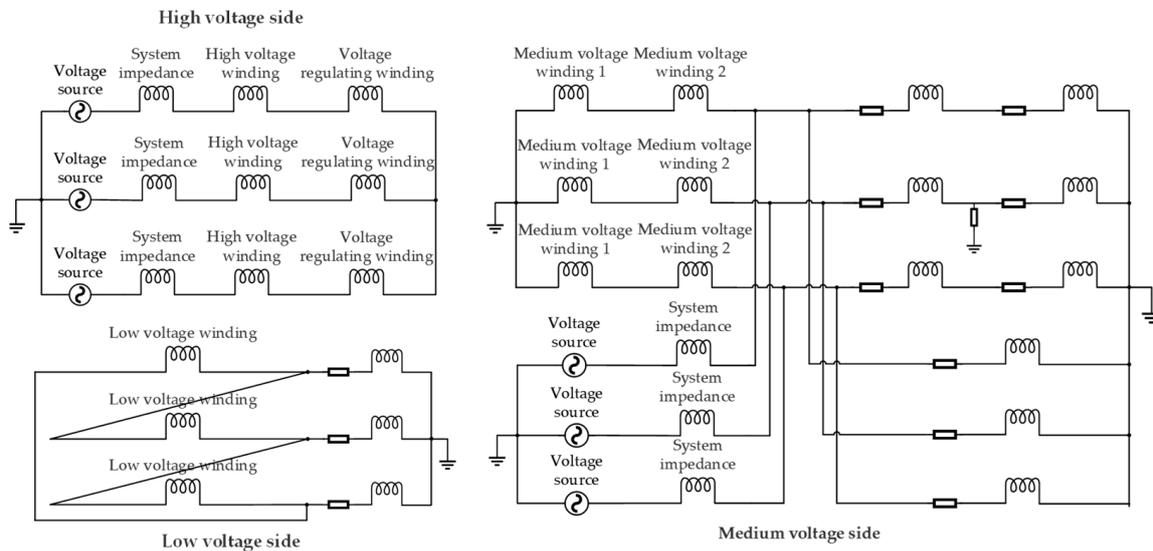


Figure 3. Diagram of the external circuit.

Table 1. Electrical parameters of transformer.

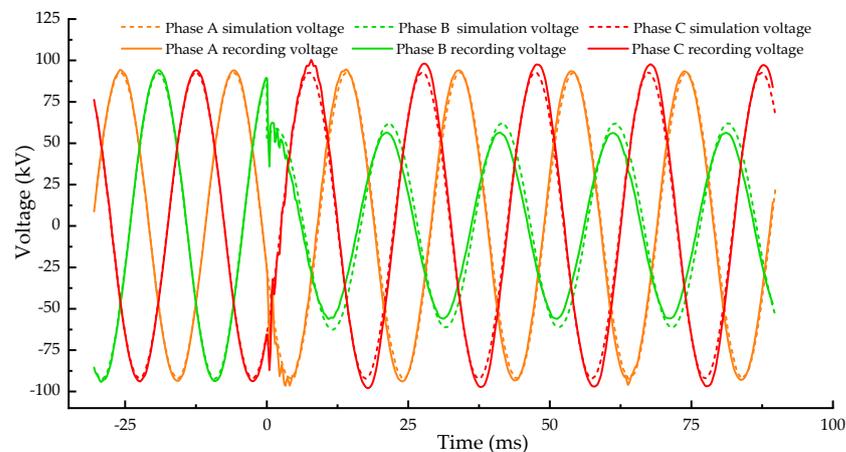
Name	Parameter Value	Name	Parameter Value
rated capacity (MVA)	240/240/80	HV winding turns	533
rated voltage (kV)	220/115/10.5	MV winding turns	280
rated current (kA)	0.63/1.2/4.4	LV winding turns	44
connection group number	YNyn0d11	VR winding turns	64

**Table 2.** Structural dimensions of transformer.

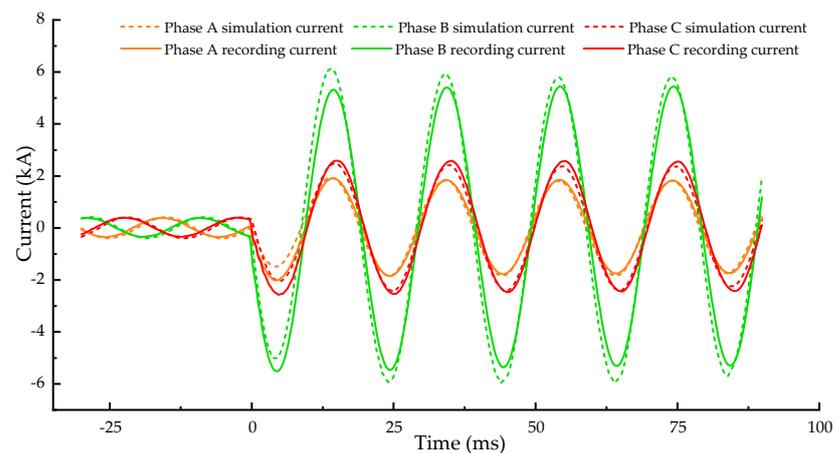
Name	Parameter Value	Name	Parameter Value
core diameter (mm)	1200	MV winding 1 diameter (mm)	1040–1085
core window height (mm)	1800	VR winding diameter (mm)	975–990
center distance (mm)	2300	HV winding diameter (mm)	825–925
yoke height (mm)	1200	MV winding 2 diameter (mm)	730–775
winding height (mm)	1600	LV winding diameter (mm)	650–680

3.2. Verification of Transformer Model

To verify the correctness of the model, the B-phase grounding fault occurring on the medium-voltage side is compared with the fault recording data. Figures 4 and 5 show the comparison of the three-phase voltage and current of the medium-voltage side between the simulated results and fault recording data.



**Figure 4.** Comparison of three-phase voltage between simulated results and fault recording data.



**Figure 5.** Comparison of three-phase current between simulated results and fault recording data.

Figure 4 shows a noticeable reduction in the peak voltage of the fault phase and little change in the peak voltages of the other two phases. The simulation results of the medium-side voltage are highly consistent with the fault recording data. The simulation results reveal significant increases in the peak values of the three-phase currents (Figure 5). In particular, the peak current of phase B increased the most, reaching 5.8 kA, which was 14 times higher than before the fault. In addition, the peak currents of phase A and phase C reached 1.81 kA and 2.37 kA, which were 4.8 times and 6.6 times those before the fault, respectively.

The simulation results of three-phase currents also match well with the actual fault recording data, and a detailed comparison is shown in Table 3. The current simulation data before the fault are highly consistent with the recorded wave data, with an error of only 1.5%. After the fault, the current simulation data are slightly different from the recording data, with a maximum error of 7.4%.

**Table 3.** Comparison of current simulation results and fault recording data.

	Parameter	Simulation Results (kA)	Fault Recording Data (kA)	Error%
before fault	phase A current	0.393	0.387	1.5
	phase B current	0.393	0.387	1.5
	phase C current	0.393	0.387	1.5
after fault	phase A current	1.81	1.84	1.6
	phase B current	5.8	5.45	6.4
	phase C current	2.37	2.56	7.4

The good agreement between the above simulation results and the fault recording data shows that the simulation model is accurate and feasible in simulating the external short-circuit fault of the power transformer. Therefore, this simulation model is utilized to simulate various external short-circuit fault conditions and generate current data as original data for collection. In this study, typical external transformer short-circuit faults are simulated, encompassing single-phase grounding fault (single-phase GF), two-phase short-circuit fault (two-phase SCF), two-phase grounding fault (two-phase GF), and three-phase grounding fault (three-phase GF), yielding 210 sets of current data for transformer external short-circuit faults. In the subsequent analysis, 150 sets of data are used as training samples and 60 sets of data serve as testing samples. The data set distribution is shown in Table 4.

**Table 4.** Data set distribution.

Fault Type	Number of Total Fault Data	Number of Testing Fault Data	Number of Training Fault Data
HV-side single-phase GF	21	6	15
HV-side two-phase SCF	21	6	15
HV-side two-phase GF	21	6	15
HV-side three-phase GF	7	2	5
MV-side single-phase GF	21	6	15
MV-side two-phase SCF	21	6	15
MV-side two-phase GF	21	6	15
MV-side three-phase GF	7	2	5
LV-side single-phase GF	21	6	15
LV-side two-phase SCF	21	6	15
LV-side two-phase GF	21	6	15
LV-side three-phase GF	7	2	5

## 4. The Algorithm Flow and Impact of Kernel Functions on Identification Accuracy

### 4.1. Algorithm Flow of Transformer Fault Identification

The flowchart in Figure 6 illustrates the process of transformer fault identification based on SVM. This process mainly consists of five steps:

- (1) Extract features from all obtained fault data as training and testing samples. For each short-circuit fault, the total variation rate of the peak currents on the HV side, MV side, and LV side of the three phases are extracted separately, resulting in nine features and forming a nine-dimensional feature vector. This feature vector corresponds to  $\varphi(x)$  in Equation (1).

- (2) Use all training samples to train the fault-side classification model to establish high-, medium-, and low-voltage fault classification models, which can determine the fault side. For these three categories, labels 1, 2, and 3 are assigned, respectively, corresponding to  $y_i$  in Equations (3)–(5).
- (3) Use HV-, MV-, and LV-side fault samples to train fault-type classification models, respectively. Taking the MV side as an example, the MV fault-type classification model can classify the input samples into MV single-phase GF, MV two-phase SCF, MV two-phase GF, and MV three-phase GF.
- (4) Use single-phase GF samples, two-phase SCF samples, and two-phase GF samples to train fault-phase classification models, respectively. These models can determine the fault phase.
- (5) Input the external short-circuit testing samples into the fault-side, fault-type, and fault-phase classification models in sequence to obtain the fault side, the fault type, and the fault phase.

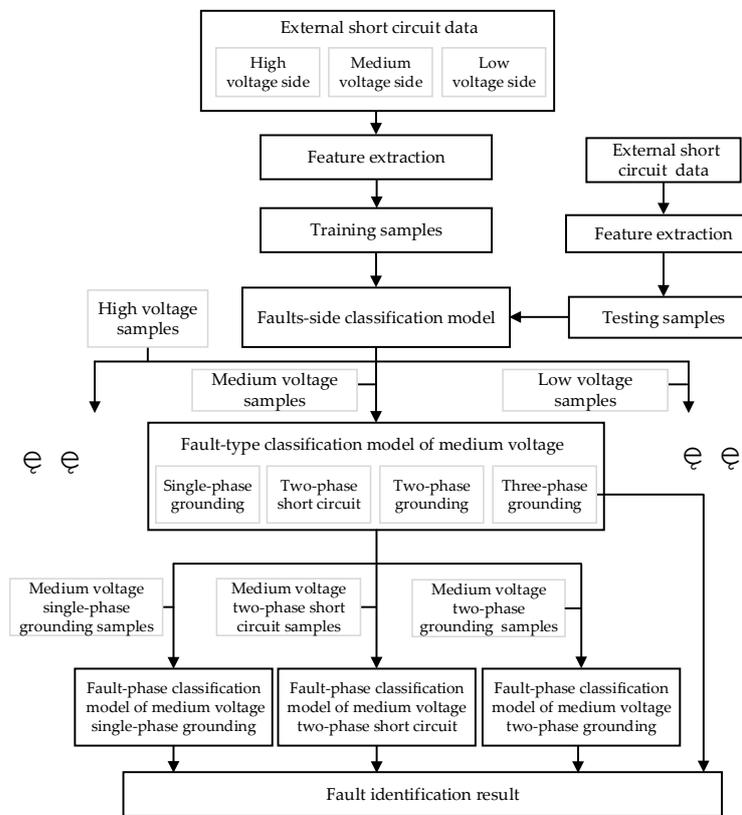


Figure 6. Algorithm flow of fault identification.

In this paper, nine characteristic features are defined for each fault current data, and the specific definition methods are as follows:

$$\begin{aligned}
 HA &= \frac{|I_{HA-} - I_{HA+}|}{I_{HA}} + \frac{|I_{HA-} - I_{HA-}|}{I_{HA}} \\
 HB &= \frac{|I_{HB-} - I_{HB+}|}{I_{HB}} + \frac{|I_{HB-} - I_{HB-}|}{I_{HB}} \\
 HC &= \frac{|I_{HC-} - I_{HC+}|}{I_{HC}} + \frac{|I_{HC-} - I_{HC-}|}{I_{HC}}
 \end{aligned} \tag{10}$$

$$\begin{aligned}
 MA &= \frac{|I_{MA-} - I_{MA+}|}{I_{MA}} + \frac{|I_{MA-} - I_{MA-}|}{I_{MA}} \\
 MB &= \frac{|I_{MB-} - I_{MB+}|}{I_{MB}} + \frac{|I_{MB-} - I_{MB-}|}{I_{MB}} \\
 MC &= \frac{|I_{MC-} - I_{MC+}|}{I_{MC}} + \frac{|I_{MC-} - I_{MC-}|}{I_{MC}}
 \end{aligned} \tag{11}$$

$$\begin{aligned}
 LA &= \frac{|I_{LA-} - I_{LA+}|}{I_{LA}} + \frac{|I_{LA-} - I_{LA-}|}{I_{LA}} \\
 LB &= \frac{|I_{LB-} - I_{LB+}|}{I_{LB}} + \frac{|I_{LB-} - I_{LB-}|}{I_{LB}} \\
 LC &= \frac{|I_{LC-} - I_{LC+}|}{I_{LC}} + \frac{|I_{LC-} - I_{LC-}|}{I_{LC}}
 \end{aligned}
 \tag{12}$$

In the definition of the above characteristic features, *HA*, *HB*, and *HC* represent the total variation rate of the three-phase current peak values on the HV side; *MA*, *MB*, and *MC* indicate the total variation rate of the three-phase current peak values on the MV side, respectively; and *LA*, *LB*, and *LC* represent the total variation rate of the three-phase current peak values on the LV side.

$I_{HA}$  represents the peak current of phase A on the HV side before the fault.  $I_{HA+}$  and  $I_{HA-}$  indicate the first positive and negative half-cycle peak currents on the HV side after the fault, respectively.  $I_{HB}$ ,  $I_{HB+}$ ,  $I_{HB-}$  and  $I_{HC}$ ,  $I_{HC+}$ ,  $I_{HC-}$  represent the corresponding variables of phase B and phase C on the HV side, respectively.  $I_{MA}$ ,  $I_{MA+}$ ,  $I_{MA-}$ ,  $I_{MB}$ ,  $I_{MB+}$ ,  $I_{MB-}$ ,  $I_{MC}$ ,  $I_{MC+}$ ,  $I_{MC-}$  and  $I_{LA}$ ,  $I_{LA+}$ ,  $I_{LA-}$ ,  $I_{LB}$ ,  $I_{LB+}$ ,  $I_{LB-}$ ,  $I_{LC}$ ,  $I_{LC+}$ ,  $I_{LC-}$  represent the corresponding parameters of the MV side and the LV side, respectively.

#### 4.2. The Impact of Kernel Functions on Identification Accuracy

The SVM algorithm employs kernel functions to map input data into a higher-dimensional space, making them linearly separable. Different kernel functions handle data mapping in various ways, so the chosen kernel function should take into account the attributes of the data and the nature of the problem. In this section, we analyze the identification accuracy of four different kernel functions, including the LN, the PL, the RBF, and the SIG, whose definitions are described in Section 2.1.

A total of 30 testing samples were used to discuss the influence of kernel functions. Figure 7 depicts the influence of the four different kernel functions on identification accuracy. The identification accuracies of the LN, the PL, and the RBF all reach 100% in fault-side identification, while the performance of the SIG is relatively worse. In fault-type identification, the PL and the RBF perform well, achieving an identification accuracy of 96.7%, while the other two kernel functions are slightly inferior. When identifying the fault phase, all four kernel functions achieve a 96.7% identification accuracy. In general, the identification accuracies of the PL and the RBF are higher than those of the other two kernel functions. Due to the outstanding generalization ability and fewer parameter requirements, the radial basis function kernel is adopted in this paper.

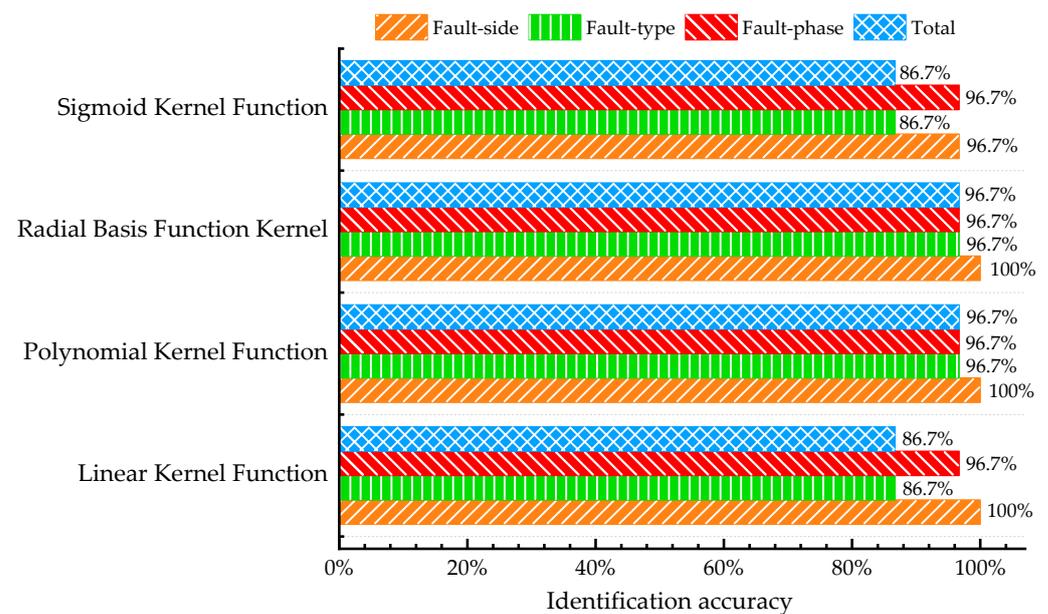


Figure 7. Influence of identification accuracy under different kernel functions.

### 5. Identification Results

#### 5.1. Identification Results Based on Simulation Data

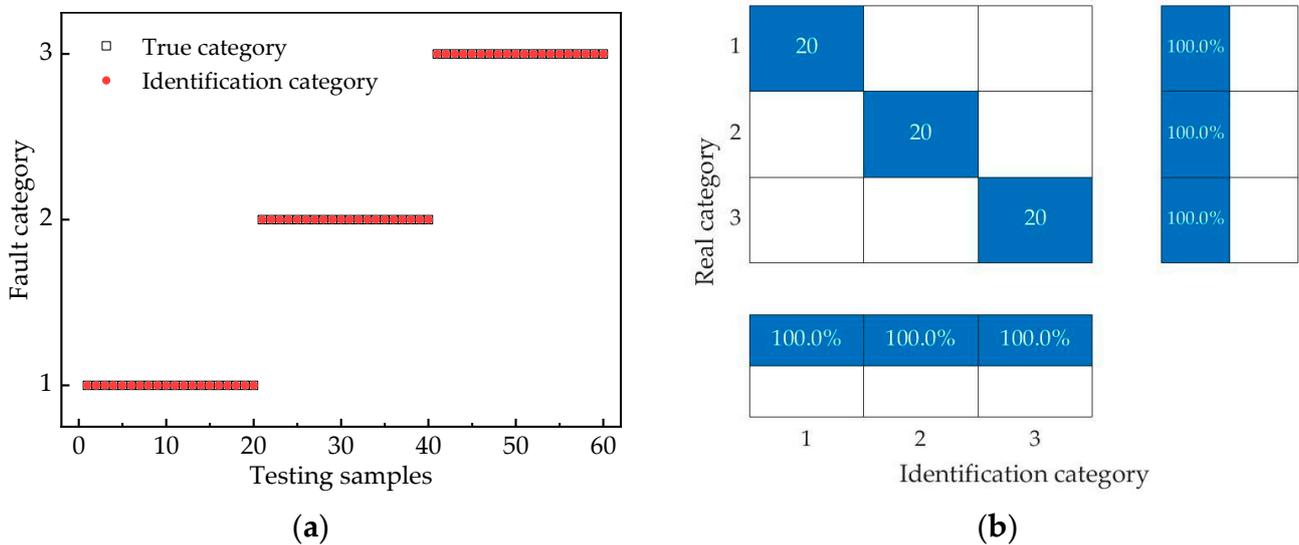
##### 5.1.1. Fault-Side Identification Results

According to the fault identification flow shown in Figure 6, fault-side identification is carried out first. The distribution of testing samples is presented in Table 5.

**Table 5.** The distribution of testing samples for fault-side identification.

Sample Labels	1	2	3
fault side	HV side	MV side	LV side
number of testing samples	20	20	20

The fault-side identification results and confusion matrix are displayed in Figure 8. It can be observed that the identification accuracy can reach 100% when determining the fault side on the HV side, MV side, and LV side.



**Figure 8.** Fault-side identification results and confusion matrix. (a) Identification results; (b) Confusion matrix.

##### 5.1.2. Fault-Type Identification Results

The second step is to identify the fault type. The distribution of testing samples for fault-type identification is displayed in Table 6, with the identification results depicted in Figure 9. The identification accuracy for fault types on the HV side is 95%, and the identification accuracy of the MV and LV side can reach 100%.

**Table 6.** The distribution of testing samples for fault-type identification.

Sample Labels	1	2	3	4
fault type	Single-phase GF	Two-phase SCF	Two-phase GF	Three-phase GF
number of testing samples of HV side	6	6	6	2
number of testing samples of MV side	6	6	6	2
number of testing samples of LV side	6	6	6	2

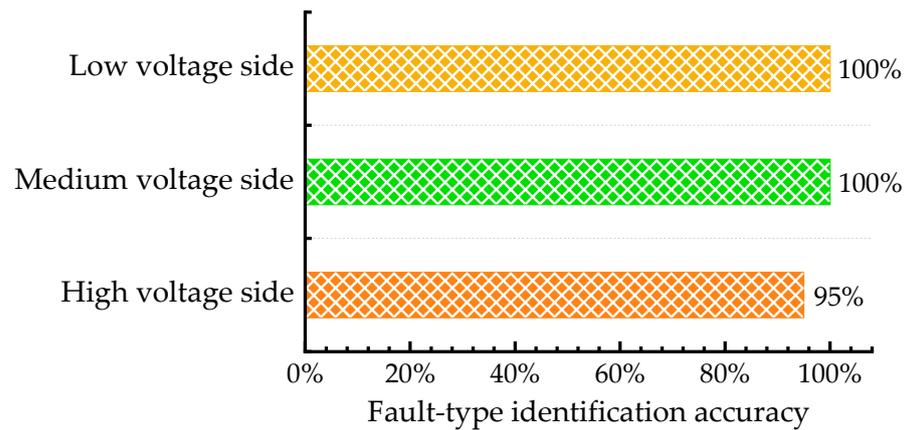


Figure 9. Fault-type identification accuracy.

For the high-voltage side, the detailed fault-type identification results and confusion matrix are shown in Figure 10. The identification performance is good when recognizing single-phase GF, two-phase GF, and three-phase GF types. However, during the identification of two-phase SCF, one sample is misclassified as a single-phase GF.

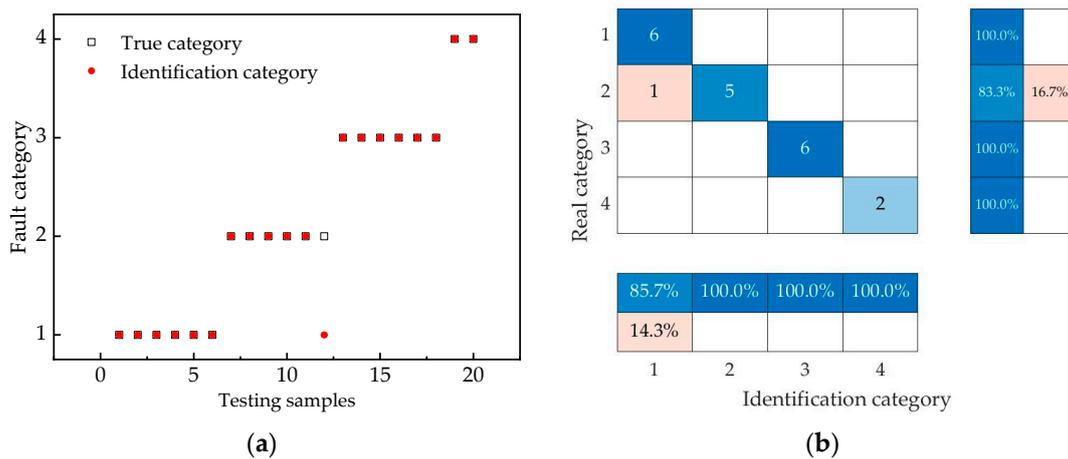


Figure 10. Fault-type identification results and confusion matrix of high-voltage side. (a) Identification results; (b) Confusion matrix.

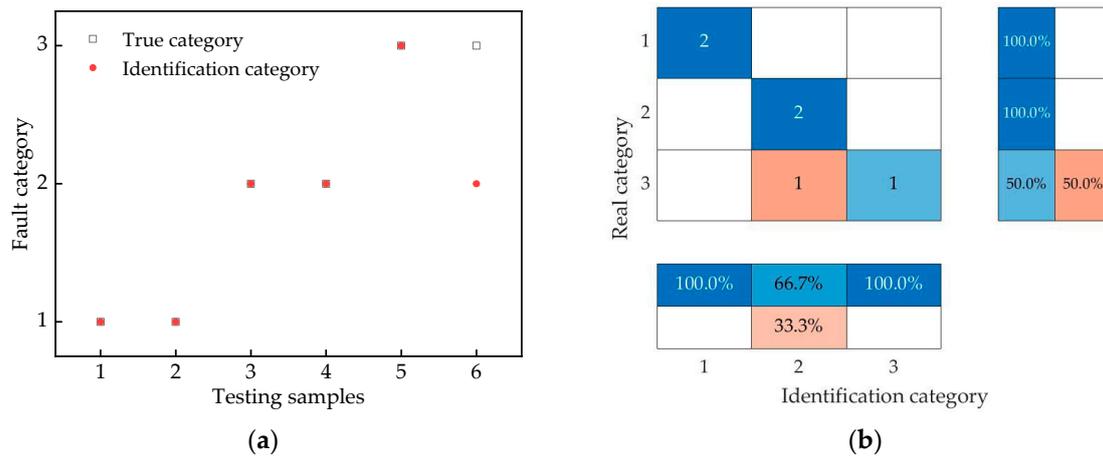
### 5.1.3. Fault-Phase Identification Results

The third step is to identify the fault phase. Taking the two-phase SCF of the HV side as an example, Table 7 shows the distribution of testing samples for fault-phase identification.

Table 7. The distribution of testing samples for fault-phase identification.

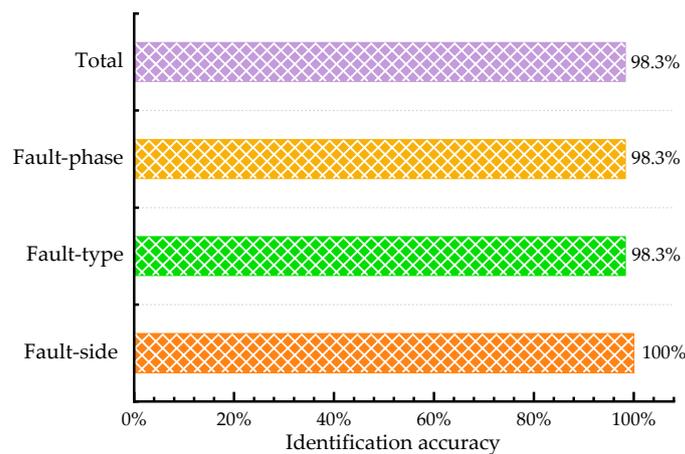
Sample Labels	1	2	3
fault phase	A/B Two-phase SCF	A/C Two-phase SCF	B/C Two-phase SCF
number of testing samples	2	2	2

The detailed fault-phase identification results and confusion matrix are shown in Figure 11. Only one of the six samples has an incorrect identification result: one of the samples with a B/C two-phase SCF is judged as an A/C two-phase SCF.



**Figure 11.** Fault-phase identification results and confusion matrix of two-phase short circuit of high-voltage side. (a) Identification results; (b) Confusion matrix.

Figure 12 illustrates the fault-side, fault-type, and fault-phase identification accuracies of all tested samples. One can see that the fault-side identification accuracy is 100% and the fault-type and fault-phase identification accuracies are both 98.3%. The total identification accuracy reaches 98.3%, verifying the effectiveness of the SVM-based fault identification method proposed in this paper.



**Figure 12.** Identification accuracy.

### 5.2. Comparison of Different Algorithms

To validate the applicability of the SVM algorithm in this research, random forest (RF) and decision tree (DT) algorithms are employed to identify the external faults of transformers, and the results are compared with those of SVM.

Figure 13 illustrates the accuracy of the three different algorithms in identifying the fault side, fault type, and fault phase. SVM exhibits superior performance in fault-side and fault-phase identification, achieving accuracies of 100% and 98.3%, respectively, surpassing RF and DT. In fault-type identification, all algorithms demonstrate relatively close accuracies, with SVM and RF slightly outperforming DT. These results suggest that SVM may have advantages in fault identification.

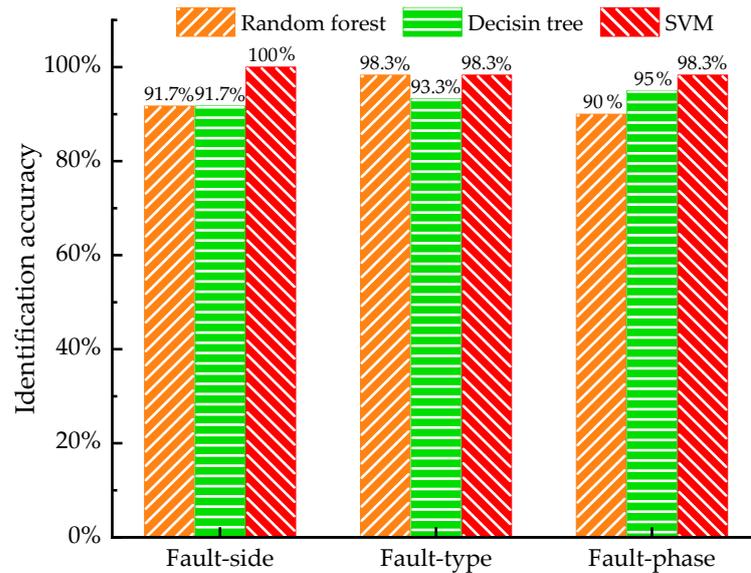


Figure 13. Identification accuracy of different algorithms.

5.3. Identification Results Based on Fault Recording Data

To further validate the efficacy of the identification method proposed in this study, an examination based on fault recording data was conducted. In this case, the fault recording data originated from an actual external short-circuit transformer fault that occurred in a substation. Specifically, the fault is a B-phase GF on the medium-voltage side. The actual recorded current waveform for this fault is shown in Figure 14. Table 8 shows the current characteristic features extracted from fault recording data according to the method described in Section 4.1. Utilizing the method proposed in this paper, the accurate determination of the fault side, fault type, and fault phase is achievable, further validating the effectiveness of this method.

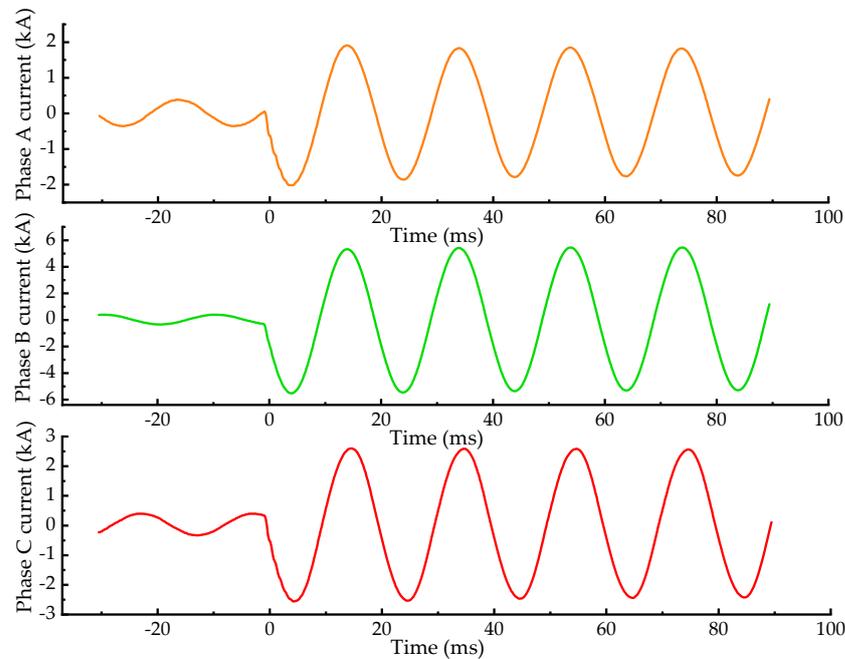


Figure 14. Recorded current waveform of B-phase GF on the MV side.

**Table 8.** Current characteristic features of fault recording data.

Name	HA	HB	HC	MA	MB	MC	LA	LB	LC
parameter values	0.606	12.911	2.176	8.155	28.034	0.606	12.911	2.176	8.155

The above analysis demonstrates that the method proposed in this paper can accurately identify the fault side, fault type, and fault phase of external transformer faults. Therefore, this method has significance in transformer fault identification and can be widely applied in practical engineering.

## 6. Conclusions

To accurately identify the external short-circuit faults of transformers, we propose a fault identification method based on SVM. By extracting features from fault currents, three classification models are utilized to precisely identify the fault side, fault type, and fault phase. The main conclusions are as follows:

- (1) A method for extracting features from external short-circuit fault currents is defined, comprising nine features, namely the total variation rates of three-phase peak currents for the HV, MV, and LV sides. The results indicate that this method of defining features is suitable for identifying external short-circuit faults in transformers.
- (2) The influence of four different kernel functions on the accuracy of SVM classification models is discussed, including the LN, the PL, the RBF, and the SIG. The identification accuracy of the RBF is higher than the other kernel functions.
- (3) The identification results for external transformer faults are analyzed using 60 simulation data as testing samples, with an identification accuracy of up to 98.3%. Furthermore, the classification model is validated by using a set of actual fault current recording data, where the fault side, fault type, and fault phase can be identified correctly. The method proposed in this paper was demonstrated to accurately identify the external short-circuit fault of transformers.

This research provides a feasible solution for the precise identification of external short-circuit faults, which can be widely applied in practical engineering. This method utilizes the change rate of peak short-circuit current as a characteristic feature. In future research, incorporating more features, such as the frequency, phase angle, and characteristics of zero-sequence current signals, can further enhance the accuracy of fault identification.

**Author Contributions:** Conceptualization, Z.R.; methodology, X.S., S.J. and Z.L.; writing—original draft preparation, H.D.; supervision, X.L.; project administration, L.C. and Z.M. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by Guangdong Key Laboratory of Electric Power Equipment Reliability, Electric Power Research Institute of Guangdong Power Grid Co., Ltd., Guangzhou, Guangdong 510080, China (No. GDDKY2022KF01).

**Data Availability Statement:** Data are contained within the article.

**Conflicts of Interest:** Author Linglong Cai, Zhiqin Ma, Zhangquan Rao, Xiang Shu, Shuo Jiang and Zhongxiang Li were employed by Electric Power Research Institute of Guangdong Power Grid, Guangzhou, Guangdong, China. 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.

## References

1. Wang, F.; Hu, X.; Qian, Y.; Chen, P.; Ma, X. Research and review on monitoring technology for transformer winding vibration. *Guangdong Electr. Power* **2018**, *31*, 52–61.
2. Qu, Y.; Zhao, H.; Ma, L.; Zhao, S.; Mi, Z. Multi-depth neural network synthesis method for power transformer fault identification. *Proc. CSEE* **2021**, *41*, 8223–8231.
3. Wang, Q.; Jia, S.; Liang, D.; Zhao, Y.; Bi, J.; Yang, Z. Temperature characteristics analysis and guide structure optimization of low-frequency transformer based on multi-physics field simulation. *High Volt. Eng.* **2023**. [[CrossRef](#)]

4. Xie, L.; Zhang, F.; Liu, C.; Zheng, J.; Luo, H.; Xi, R.; Ji, S.; Zhao, J. Research on vibration and noise of 110kV transformer based on staged finite element method. *High Volt. Appar.* **2023**, *59*, 22–34.
5. Zheng, Y.; Gond, X.; Pan, S.; Sun, J.; Deng, J. Analysis on leakage flux characteristics of turn-to-turn short-circuit fault for power transformer. *Autom. Electr. Power Syst.* **2022**, *46*, 121–127.
6. Zhao, B.; Yang, M.; Diao, H.R.; An, B.; Zhao, Y.C.; Zhang, Y.M. A novel approach to transformer fault diagnosis using IDM and naive credal classifier. *Int. J. Electr. Power Energy Syst.* **2019**, *105*, 846–855. [[CrossRef](#)]
7. Shang, H.; Xu, J.; Zheng, Z.; Qi, B.; Zhang, L. A novel fault diagnosis method for power transformer based on dissolved gas analysis using hypersphere multiclass support vector machine and improved D–S evidence theory. *Energies* **2019**, *12*, 4017. [[CrossRef](#)]
8. Faiz, J.; Milad, S. Assessment of computational intelligence and conventional dissolved gas analysis methods for transformer fault diagnosis. *IEEE Trans. Dielectr. Electr. Insul.* **2018**, *25*, 1798–1806. [[CrossRef](#)]
9. Zou, A.; Deng, R.; Mei, Q.; Zou, L. Fault diagnosis of a transformer based on polynomial neural networks. *Clust. Comput.* **2019**, *22*, 9941–9949. [[CrossRef](#)]
10. Liu, Y.; Song, B.; Wang, L.; Gao, J.; Xu, R. Power transformer fault diagnosis based on dissolved gas analysis by correlation coefficient-DBSCAN. *Appl. Sci.* **2020**, *10*, 4440. [[CrossRef](#)]
11. Huang, Y.; Wu, W.; Kuo, C. Application of fault overlay method and CNN in infrared image of detecting inter-urn short-circuit in dry-type transformer. *Electronics* **2023**, *12*, 181. [[CrossRef](#)]
12. Doolindachaporn, A.; Callender, G.; Lewin, P.; Simonson, E.; Wilson, G. Data driven transformer thermal model for condition monitoring. *IEEE Trans. Power Deliv.* **2022**, *37*, 3133–3141. [[CrossRef](#)]
13. Cheng, L.; Yu, T.; Wang, G.; Yang, B.; Zhou, L. Hot spot temperature and grey target theory-based dynamic modelling for reliability assessment of transformer oil-paper insulation systems: A practical case study. *Energies* **2018**, *11*, 249. [[CrossRef](#)]
14. Deng, Y.; Ruan, J.; Dong, X.; Huang, D.; Zhang, C. Inversion detection method of oil-immersed transformer abnormal heating state. *IET Electr. Power Appl.* **2023**, *17*, 134–148. [[CrossRef](#)]
15. Shiravand, V.; Faiz, J.; Samimi, M.H.; Mehrabi Kermani, M. Prediction of transformer fault in cooling system using combining advanced thermal model and thermography. *IET Gener. Transm. Distrib.* **2021**, *15*, 1972–1983. [[CrossRef](#)]
16. Soleimani, M.; Faiz, J.; Nasab, P.S.; Moallem, M. Temperature measuring-based decision-making prognostic approach in electric power transformers winding failures. *IEEE Trans. Instrum. Meas.* **2020**, *69*, 6995–7003. [[CrossRef](#)]
17. Bagheri, M.; Nezhivenko, S.; Naderi, M.S.; Zollanvari, A. A new vibration analysis approach for transformer fault prognosis over cloud environment. *Int. J. Electr. Power Energy Syst.* **2018**, *100*, 104–116. [[CrossRef](#)]
18. Duan, X.; Zhao, T.; Liu, J.; Zhang, L.; Zou, L. Analysis of winding vibration characteristics of power transformers based on the finite-element method. *Energies* **2018**, *11*, 2404. [[CrossRef](#)]
19. Jiang, P.; Zhang, Z.; Dong, Z.; Wu, Y.; Xiao, R.; Deng, J.; Pan, Z. Research on distribution characteristics of vibration signals of  $\pm 500$  kV HVDC converter transformer winding based on load test. *Int. J. Electr. Power Energy Syst.* **2021**, *132*, 107200. [[CrossRef](#)]
20. Zhao, L.; Zhang, Z.; Zhang, J.; Huang, X.; Ren, J. Diagnosis methods for transformer faults based on vibration signal under fluctuating operating conditions. *High Volt. Eng.* **2020**, *46*, 3925–3933.
21. Deng, X.; Zhang, Z.; Zhu, H.; Yan, K. Early fault diagnosis of transformer winding based on leakage magnetic field and DSAN learning method. *Front. Energy Res.* **2023**, *10*, 1058378. [[CrossRef](#)]
22. Yan, C.; Zhang, P.; Xu, Y.; Liu, S.; An, T. Bidirectional field–circuit coupling analysis of converter transformer inter-tap short-circuit faults in on-load tap changers. *IET Gener. Transm. Distrib.* **2022**, *16*, 4750–4760. [[CrossRef](#)]
23. Cabanas, M.F.; Pedrayes, F.; Melero, M.G.; Rojas, C.H.; Orcajo, G.A.; Cano, J.M.; Norniella, J.G. Insulation fault diagnosis in high voltage power transformers by means of leakage flux analysis. *Prog. Electromagn. Res.-Pier* **2011**, *114*, 211–234. [[CrossRef](#)]
24. Haghjoo, F.; Mohammadi, H. Planar sensors for online detection and region identification of turn-to-turn faults in transformers. *IEEE Sens. J.* **2017**, *17*, 5450–5459. [[CrossRef](#)]

**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.