CIT. Journal of Computing and Information Technology, Vol. 31, No. 2, June 2023, 73–89
doi: 10.20532/cit.2023.1005733

Automatic Diagnosis Method for Short Circuit Faults in Power Measurement Instruments Based on CNN

Xiuni Xu

College of New Energy, Longdong University, Qingyang, Gansu, China
Tianjin University of Technology and Education, Tianjin, China

The accuracy of short-circuit fault diagnosis methods for power measuring instruments is still low, which can directly lead to abnormal power data statistics. To address this issue, a convolutional neural network is used to construct an automatic diagnosis model for short circuit faults in power measuring instruments. The model utilizes wavelet packet energy spectrum to extract crucial features from the signal. The model identifies and determines the fault type based on the obtained operation characteristic data and the set short-circuit fault diagnosis criteria. The Sparrow optimization algorithm is used to optimize the model's parameters and enhance its performance. Experimental analysis revealed that after the signal features were extracted by using the wavelet energy spectrum distribution, the error value fluctuated in the range of 0.00 ~ 0.04 with regard to the measured value and was mainly around 0.02. The extracted features achieved high accuracy levels. The designed model exhibited an average diagnostic accuracy of 97.03%, surpassing the other three models by 9.24%, 7.10%, and 4.25%. The presented model can improve the precision and productivity of fault detection, support the safety and reliability of power system operations, and facilitate the collection and analysis of power usage data.

ACM CCS (2012) Classification: Computing methodologies → Machine learning → Machine learning algorithms

Keywords: CNN, Power measurement instruments, Short circuit faults, Automatic diagnosis, Sparrow search algorithm

1. Introduction

Power measurement instruments are front-end acquisition components of digital power grids, providing parameter data for power parameter measurement, power quality monitoring, electrical equipment control, and other tasks. Therefore, they are widely used in various power control systems. However, due to factors such as the surrounding magnetic field environment and equipment aging, short circuit faults often occur in power measurement instruments [1–2]. The faults in power measurement instruments directly lead to abnormal statistics of power data, and the accuracy of statistical data directly affects the decision-making of the power department regarding power grid allocation. Therefore, automatic diagnosis of faults in power measurement instruments is of great importance.

Short circuit faults are one of the most common types of faults in power measurement instruments [3–4]. Existing fault diagnosis methods mainly focus on qualitative analysis, which is not highly accurate. The commonly used method for quantitative analysis is fault tree analysis, which is intuitive, easy to understand, and can achieve fault analysis and diagnosis. However, as time goes by, the fault location also changes continuously, and this method determines the probability value of the fault cause as a constant, which does not match the actual situation [5].

Traditional methods for fault diagnosis rely heavily on manual detection and qualitative analysis, which can be time-consuming, costly,

and prone to human error. Automatic and accurate diagnosis is not achieved through these methods. To address these issues, new methods for fault diagnosis are needed. This can result in misjudgments or missed detections. The measurement technology currently in use has limited scalability for diagnosing faults in measuring instruments. As the power grid expands and power equipment increases, many fault diagnosis technologies struggle to achieve ideal diagnostic accuracy and efficiency, and face challenges in operational stability.

To address the aforementioned issues, we have developed a new automated diagnostic model for detecting short circuit faults in power measuring instruments. This model incorporates both convolutional neural network (CNN) and long and short-term memory network (LSTM) approaches. The objective is to enable real-time monitoring of power measuring instrument failures and ensure the delivery of high-quality power supply services. The research is innovative in two aspects. Firstly, artificial intelligence technology is applied to short-circuit fault diagnosis of power measuring instruments, which improves the accuracy of statistical data of the power sector. Secondly, the Sparrow search algorithm is optimized by using Logistic chaotic mapping, and the parameters of the model are optimized by using the optimized algorithm, which further improves the diagnostic accuracy of the model.

The research makes two technical contributions. Firstly, it successfully constructs an automatic diagnosis model for short circuit faults in power measuring instruments based on CNN and LSTM. This model enables automatic fault diagnosis of power measuring instruments. Secondly, it innovatively combines the Sparrow search algorithm with Logistic chaotic mapping to improve the diagnostic accuracy and generalization ability of the model. The automatic diagnosis method proposed can reduce the time and cost of manual detection and qualitative analysis. Additionally, the proposed model can be extended to other fields of fault detection, providing a new reference for related areas.

The paper is organized into 3 parts. The first part contains the literature review, which analyzes the current research status of the field at home and abroad and discusses the limitations and application of existing methods for diagnosing short circuit faults in power measurement instruments. The second part outlines the method statement, which provides a detailed description of the technologies used in the research and the methods for technical improvement in constructing the model. The third part presents the experimental analysis, which analyzes the performance of the constructed model.

2. Related Work

Scholars both domestically and internationally have conducted extensive academic research on the diagnosis and analysis of short circuit faults. Alloui *et al.* proposed an effective online diagnostic method to improve the reliability and safety of industrial facilities by detecting point-to-point short circuit faults. This method was based on the computation and detection of the ratio between zero voltage and positive voltage symmetrical components. The results showed that this method still maintains high reliability, speed, and accuracy in diagnosing short circuit faults [6].

Khelif *et al.* discussed the short circuit faults of branch capacitors in variable frequency asynchronous motors to improve the performance of electromechanical systems. They provided simulation results for this type of fault and analyzed and explained the impact on rectifiers, inverters, and asynchronous motor performance [7]. Wang *et al.* aimed to evaluate the fault characteristics of different battery packs accurately and reliably by calculating the correlation coefficient between adjacent battery voltages. Then, they trained a principal component analysis model using all the correlation coefficient signals. Based on this, they designed thresholds based on kernel density estimation and comprehensive statistics to monitor all CC signals at each sampling moment and determine if a fault has occurred [8].

For stator windings, Zhang *et al.* proposed a new diagnostic method for inter-turn short circuit (SC) faults to achieve accurate and efficient diagnosis. They also analyzed in detail the influence of inter-turn fault positions on motor magnetic flux and motor parameters through finite element analysis. The experiment outcomes showed that the accuracy of fault detection using this method is above 90% [9]. Zhang *et al.* invented a fault diagnosis strategy for ensuring the safety and reliable operation of electric vehicles by detecting and evaluating internal soft

short circuit faults in lithium-ion battery packs. This method was based on incremental capacity analysis, which extracts fault features from the data to make them more easily distinguishable from small voltage differences. Then, the local outlier factor method was used to calculate the uniformly distributed outlier factor values of each individual cell in the battery pack to detect short circuit faults [10].

CNN is a deep neural network that trains a sample set using explicit labels. Under the current developing stage of deep learning technology, its application has become increasingly widespread. Xu *et al.* used the CNN algorithm in deep learning to automatically extract activity features related to human life, aiming to achieve human activity recognition that differs from traditional feature extraction methods. They optimized the CNN parameters using the stochastic gradient descent algorithm to solve the problem of human activity recognition [11].

Wieczorek *et al.* proposed a face detection model in dangerous scenes for rescue teams' aid and, thus, speeding up the search for people in need of help. The model used a lightweight CNN framework for face detection in dangerous situations such as mines and avalanches. The experiment outcomes showed that the model detection accuracy is above 99% and the precision is above 98% [12].

Paraskevopoulos *et al.* addressed the problem of gradual deterioration in three-phase induction motors in many use cases. They obtained signal data by simulating asynchronous motors. They invented a wavelet-CNN for detecting systematic faults. The experiment outcomes showed that this method can identify more difficult and complex obstacles compared to the use of the naked eye [13].

Pan and colleagues developed DeepOPF, a deep CNN approach, to efficiently and reliably solve the security-constrained direct current optimal power flow problem. The authors first trained a CNN to identify the mapping between loads and generations. They then reconstructed the phase angle directly using power flow equations from the generation and load data [14].

Jiang *et al.* proposed a fault diagnosis method for infrared sleeve images by combining image segmentation with deep learning. They first built a target detection system based on the region CNN architecture to accurately extract the sleeve framework. Then, they proposed a pulse-coupled neural network based on simple linear iterative clustering to improve the region segmentation performance. The experiment outcomes show that it achieves a 98% detection accuracy [15].

Numerous methods exist for diagnosing short-circuit faults in power measurement devices based on literature. Among the most utilized methods is fault tree analysis. However, fault location is continually changing over time, and this method ascertains the probability value of the fault cause as a constant quantity, which does not align with reality. The use of artificial intelligence technology has rapidly increased, with Bayes network, Petri net, rough set, and artificial neural network being commonly used. Although Bayesian networks have advantages in fault diagnosis, they face difficulties in modeling complex power grids, resulting in complicated diagnostic models and poor scalability. Petri nets can be used for identifying faults in transmission lines. However, their fault tolerance is limited when new fault information is uncertain. Rough set theory has a significant advantage in uncertain problems. However, this method requires a high level of redundancy in fault information, making it difficult to construct a large fault diagnosis model for power measuring instruments. Therefore, this article presents a new automated diagnostic model for detecting short-circuit faults in power measurement devices, which incorporates CNN and LSTM. The proposed model aims to improve the accuracy and effectiveness of fault diagnosis, indirectly enabling safe power system operation and power data analysis.

3. Research Method

The study proposes an automated diagnostic model for power meter measurement based on the sparrow optimization CNN-LSTM. Firstly, the fault sample data is input into the network model, and the data preprocessing is performed by extracting the wave head and normalizing the data. Then, the data are categorized into training and test sets. The parameters of the network structure are set using logistic mapping and sparrow optimization algorithm. Finally, the model determines the fault type and location. The following section describes the specific construction process of the model.

3.1. Real Time Data Collection of Power Measurement Instruments Based on Clustering Fusion

Based on the location of the SC in power meter measurement, it can be divided into three types: single-phase, two-phase, and three-phase SC. The principle of three-phase SC and the equivalent circuit of power meter measurement are shown in Figure 1.

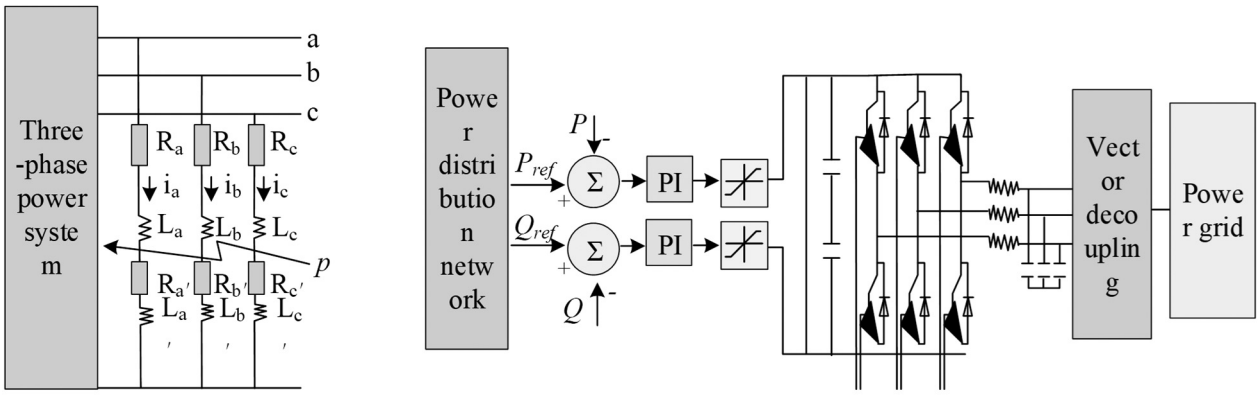
In Figure 1 (a), R_a, R_b, R_c, R'_a, R'_b, R'_c are the resistance in the line, L_a, L_b, L_c, L'_a, L'_b, L'_c are the corresponding reactance lengths, i_a, i_b, i_c represent the current on the three lines, and a, b, c represent the three circuits in the figure. In Figure 1 (b), P is the active power, Q is the reactive power, P_{ref} and Q_{ref} are the given active power reference value and reactive power reference value respectively, and L represents the power supply. Before a three-phase SC fault occurs, the internal operating circuit of the power meter measurement is always in a steady state. When a three-phase SC appears at point P in Figure 1, the circuit will be divided into two independent circuits: The upper circuit directly connected to the three-phase power system, and the power from the system will still flow into this circuit and concentrate at point P, causing a rapid increase in current and a decrease in voltage in the upper circuit. For the lower circuit, the current and voltage will decrease from their initial values at the moment of the SC to zero. The expression of the total current at the moment of the SC is shown in Equation (1).

$$i_{sc} = \sqrt{2} I_{pm} \frac{U_N}{x_s} \left| \cos(\theta - \varphi) e^{-\frac{t}{T_s}} - (1-\sigma)\cos(\theta - \varphi) e^{-\frac{t}{T_\tau}} \right| \tag{1}$$

In Equation (1), θ represents the power source voltage initial phase angle, U_N represents the rated phase voltage of the power system, x_s represents the transient reactance of the power system, σ represents the total leakage reactance coefficient, T_s represents the transient time constant of the upper circuit, T_τ represents the transient time constant of the lower circuit, I_{pm} is the periodic component amplitude of the SC current, and φ means the SC current periodic component amplitude and impedance angle. The calculation methods for I_{pm} and φ are shown in Equation (2).

$$\begin{cases} I_{pm} = \dfrac{Q}{\sqrt{R^2 + (\omega L)^2}} \\ \varphi = \arctan\left(\dfrac{\omega L}{R}\right) \end{cases} \tag{2}$$

In Equation (2), Q represents the total charge flow in the circuit, R represents the resistance value, and L represents the reactance length.



(a) Three-phase fault principle

(b) Equivalent circuit of power measuring instrument

Figure 1. Generation principle of three-phase SC and equivalent circuit of power measuring instrument.

Based on this, the principle of the short circuit fault in power meter measurement and the current and voltage characteristics are obtained. Combined with the measurement error of the meter, the criteria for fault diagnosis are set as shown in Figure 2.

According to this criteria, the data can be used for determining the existence of SC fault in the power meter measurement. Before conducting fault detection, real-time operating data of the power meter measurement needs to be collected automatically. The study selected a running signal acquisition card connected to the power meter measurement, initialized the driver program, and set the parameters for the acquisition card and A/D conversion [16]. After completing the initialization process, the acquisition signal for the meter's operating data is executed, and the results are output. In order to achieve automated data acquisition, the study designed an automatic acquisition interval of 0.5s for the acquisition card, with a continuous acquisition time of 12 hours. To achieve automatic acquisition of signal data, the acquisition card's automatic acquisition time must be set.

In power systems, the data from power measuring instruments changes rapidly. If the acquisition interval is too long, it may cause a delay in data acquisition. Therefore, the automatic collection interval is set to 0.5 seconds, taking into account the actual situation. It is important to note that data collection from power measuring instruments requires system resources. To prevent excessive consumption of system resources, the study has set the continuous collection time to 12h. The above settings enable the rational use of system resources and improve the efficiency and accuracy of data acquisition while ensuring real-time data acquisition. Additionally, this setting facilitates power measuring instrument fault detection and diagnosis, enhancing the stability and reliability of the power system. The collected operating data of the power meter measurement includes the operating current and voltage, which are sampled using digital signals. The sampling principle is shown in Equation (3).

$$\begin{cases} U = \sqrt{\dfrac{1}{N}\sum_{n=0}^{N-1} u^2(n)} \\ I = \sqrt{\dfrac{1}{N}\sum_{n=0}^{N-1} i^2(n)} \end{cases} \tag{3}$$

In Equation (3), N represents the number of samples per cycle, $u(n)$ represents the voltage values of the various lines of the power meter measurement, and $i(n)$ represents the current values of the various lines of the power meter measurement. The real-time output results of the power meter measurement collected by the acquisition device will be clustered and fused to obtain the final acquisition results. Based on the above content, the study has achieved automated data acquisition of real-time operating data for power meter measurement.

	One-way earth short circuit	Two-phase ground short circuit	Two-phase short circuit	Three-phase short circuit
Voltage characteristic	$Uc=0$	$Ub=Uc=0$	$Ua=Ub=0$	$Ua=Ub=Uc$
Current characteristic	$Ia=Ib=0$	$Ia=0$	$Ia+Ib=0$	$Ia+Ib+Ic=0$
Counting error characteristic	$0<\lambda<0.3$	$0.3<\lambda<0.5$	$0.5<\lambda<1.0$	$\lambda>1.0$

Figure 2. Setting fault diagnosis criteria.

3.2. Training and Construction of Fault Diagnosis Model Based on CNN and Data Collection

Before conducting detection and recognition, it is necessary to extract the signal variation features from the data using artificial intelligence techniques. The study used wavelet packet energy spectrum to obtain important features of the signal. The wavelet packet technique divides the original signal into multiple mutually independent frequency bands. When a SC fault appears in the transmission line, the energy of the line will undergo significant changes due to the impact of the SC current, and the energy characteristics of each frequency band will also differ. The nature of the fault can be determined by analyzing the energy distribution of each frequency band.

The fault diagnosis model of the deep neural network requires the input sample set to be in a two-dimensional matrix format, while the fault signals obtained from the system are one-dimensional [17]. Therefore, it is necessary to convert the one-dimensional signal into a two-dimensional matrix. The representation of the two-dimensional sample set is shown in Equation (4).

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ x_{h1} & x_{h2} & \cdots & x_{hm} \end{bmatrix} \tag{4}$$

In Equation (4), h represents the number of fault features for each fault sample, and m represents the number of sampling points for each fault feature. If the original fault signal displays significant fluctuation before and after the fault, using it to train a deep neural network model can result in a biased model that underestimates smaller values. Therefore, the study normalized the collected data, and the calculation method is shown in Equation (5).

$$y_i = \left(y_{\max} - y_{\min}\right) \cdot \frac{x_i - \min\limits_{1 \le j \le h}\left(x_j\right)}{\max\limits_{1 \le j \le h}\left(x_j\right) - \min\limits_{1 \le j \le h}\left(x_j\right)} + y_{\min} \tag{5}$$

In Equation (5), y_i represents the normalized result, $y_{\max}$ and $y_{\min}$ represent the max and minimum interval values of the normalized result of the original signal, respectively. The fault diagnosis of the power meter measurement in the study first needs to identify the fault nature of the data, which mainly includes the fault location and fault type. The study used CNN for fault recognition, and the network structure for fault recognition using CNN is shown in Figure 3.

The convolutional layer achieves feature mapping by sliding a 2D convolutional kernel over the input matrix and performing inner product operations. Starting from the top-left corner of the input matrix, the shared convolutional kernel slides and performs inner product operations with the corresponding positions of the input matrix. After traversing the input matrix with a 2×2 convolutional kernel, a feature map is obtained. The extraction of fault features by the convolutional layer is influenced by various factors, so it is necessary to clarify the relationship between the size of the input matrix,

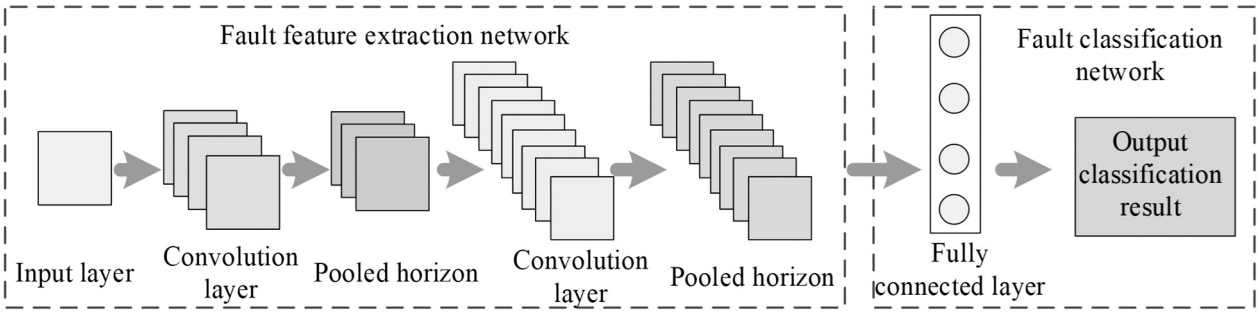


Figure 3. Network structure of CNN used for fault identification.

the convolutional kernel's size and number, the sliding stride, and the size of the output matrix [18]. The relationship is represented in Equation (6).

$$\begin{cases} h_2 = (h_1 - a)/S + 1 \\ w_2 = (w_1 - b)/S + 1 \\ d_2 = K \end{cases} \tag{6}$$

In Equation (6), $h_1 \times w_1 \times d_1$ represents the size of the input matrix, $h_2 \times w_2 \times d_2$ represents the size of the output matrix, $a \times b$ is the convolutional kernel size, K is its number, and S represents the sliding stride. The convolution operation completes the linear activation response of the input to obtain the output feature map, and further extracts fault features using non-linear activation [19]. To alleviate the problem of gradient vanishing and promote sparsity in network representation, the study utilizes the ReLU as the activation function. The calculation method of the activation function is Equation (7).

$$f(x) = \max(0, x) \tag{7}$$

Max pooling is used to denoise the data. The fault feature extraction network passes the extracted fault features from the input samples through multiple layers of convolution and pooling to the fault classification network, which then performs classification output. The study employs the softmax function to precisely classify the output of fault features in the fault classification network. This is accomplished by linearly connecting the output features and emphasizing the overall input layer information and correlation with output nodes. The specific calculation method is shown in Equation (8).

$$\begin{cases} y_i = \varphi(v_i) = \dfrac{e^{v_i}}{e^{v_1} + e^{v_2} + e^{v_3} + \ldots + e^{v_M}} \\ \varphi(v_1) + \varphi(v_2) + \ldots + \varphi(v_M) = 1 \end{cases} \tag{8}$$

In Equation (8), v_i is the weighted sum of the i-th node, and M is the number of output nodes. After feature extraction and collection, input it into the CNN model for training and recognition. The overall structure of the line fault identification method combining wavelet packet energy spectrum and CNN is shown in Figure 4.

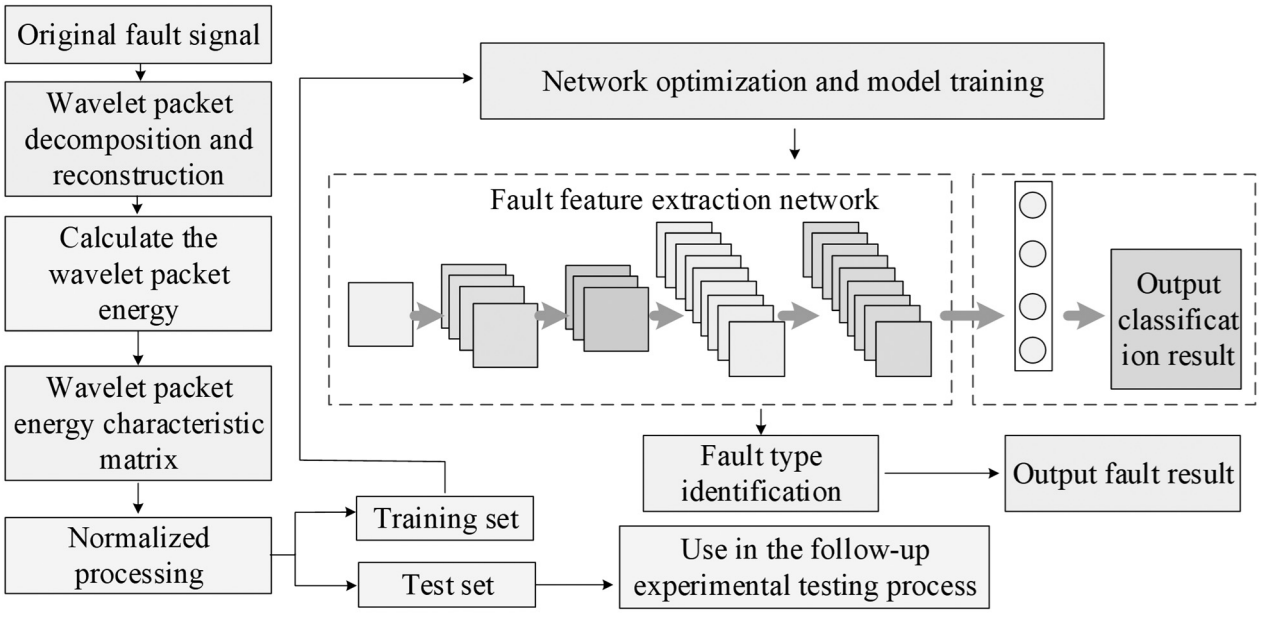


Figure 4. Overall structure of line fault identification method combining wavelet packet energy spectrum and convolutional neural network.

3.3. Design of Optimization Strategy for Fault Diagnosis Model Integrating LSTM and Sparrow Optimization Algorithm

The parameters of a deep learning network can significantly impact the final diagnostic results. Therefore, the sparrow intelligent search algorithm is utilized to optimize the model parameters. More specificaly, the study adopts a sparrow intelligence search algorithm for parameter optimization. For the individuals in the population, initial positions significantly impact optimization performance of the swarm intelligence algorithm itself. If the initial positions are too concentrated, they are prone to getting stuck in local optima and have slow convergence speed. Therefore, randomly generating initial positions cannot guarantee the stability of the optimization performance of the algorithm [19–20]. To address this, the study introduces the Logistic chaotic mapping to initialize the population distribution. The specific calculation method is shown in Equation (9).

$$x_{k+1} = \mu x_k (1 - x_k) \tag{9}$$

In Equation (9), $x \in [0, 1]$, $\mu \in [0, 4]$, $k \in Z$, μ is the chaotic parameter, and k is the number of iterations for the chaotic mapping calculation. The distribution of the sparrow population corresponds to the acquired hyperparameters. Changes in the chaotic parameter, demonstrated in Figure 5, produce changes in the sparrow population distribution through the Logistic chaotic sequence.

When the chaotic parameter is 4, the variable values generated by the Logistic mapping are mapped to the positions of the sparrow individuals, achieving population initialization. This is expressed in Equation (10).

$$x = x_{lb} + (x_{lb} - x_{ub}) x_{k+1} \tag{10}$$

In Equation (10), x_{lb} is the upper limit of the position of each individual in the sparrow population, x_{ub} is the lower limit of the position of each individual in the sparrow population, and x is the initialized position of the individual after the Logistic mapping. The real-time operational feature data extracted from the power meter is matched and compared with the set criteria for SC fault diagnosis to determine whether the current device has a short circuit fault and determine the SC fault type.

To ensure reliable functioning of the power measuring instrument, it is crucial to develop a fault location network that can precisely and effectively determine the distance of the line fault. This is important because the power measuring instrument is influenced by both internal components and timing characteristics. The CNN-LSTM fusion network combines the

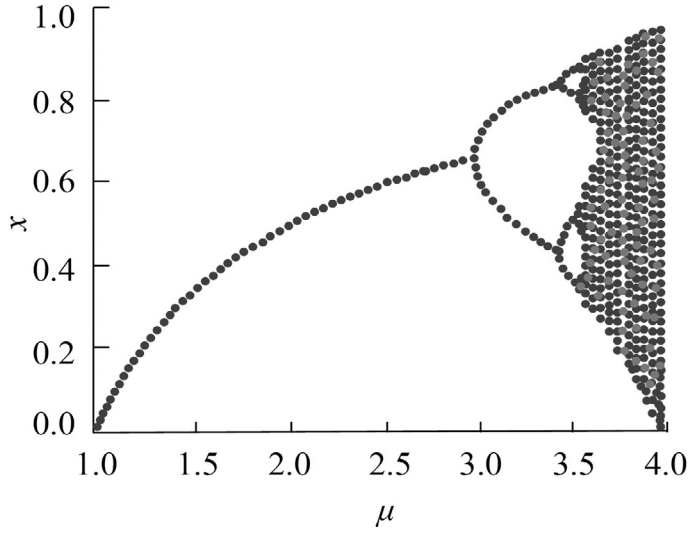


Figure 5. Distribution of Logistic chaotic sequence with changes of chaotic parameters.

distinctive features of the CNN and LSTM networks to achieve swift and precise positioning. The fault diagnosis CNN-LSTM model consists of an input layer, a CNN convolutional layer, an LSTM hidden layer, a Dropout layer, a fully connected layer, and an output layer. To prevent overfitting, random inactivation is used during dropout processing at the fully connected layer. The fault sample data is processed in the CNN convolutional layer before being passed to the LSTM hidden layer. The data passes through the Dropout and fully connected layers before reaching the output layer to determine the final fault location. The specific architecture of the model is shown in Figure 6.

The hidden layer applies the tansig transfer function, with an output value range of [−1, 1]. Equation (11) details the calculation method.

$$f(x) = \frac{2}{1+e^{-x}} - 1 \tag{11}$$

The trainlm transfer function is used in the output layer. The positions of the discoverer, entrant, and alert sparrow in the sparrow algorithm are updated through calculation. After the position update, the fitness value of individual sparrows is calculated. If the fitness value of an individual sparrow was smaller before the position update, it remains unchanged. However, if it was greater, then the sparrow population is updated and adjusted [20–21]. If the maximum number of iterations is reached, the optimal parameter is assigned to CNN-LSTM.

The root mean square error (RMSE) accurately reflects the discrepancy between the fault location result and actual fault distance and assesses the stability of the fault location effect. Therefore, the study adopts it as the loss function for the training process, with the calculation method presented in Equation (12).

$$RMSE(X, h) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(h(x_i) - y_i\right)^2} \tag{12}$$

In Equation (12), X is the set of fault samples, h is the network model referenced in fault location, n is the number of fault samples, $h(x_i)$ represents the location result of the network model for the i fault sample, and y_i is the actual location result of the i fault sample. The optimized parameter settings are shown in Table 1.

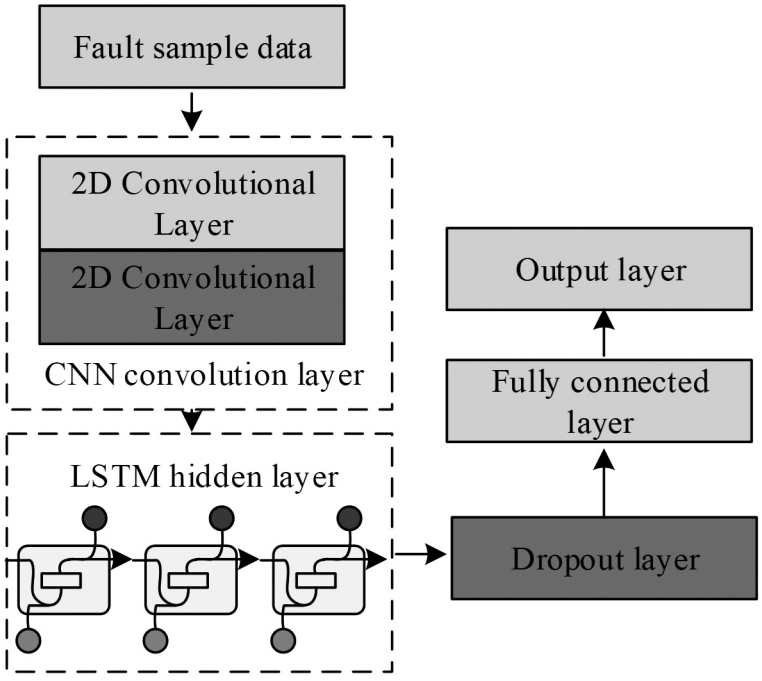


Figure 6. Architecture of the model.

Table 1. Model parameter settings optimized by the sparrow search algorithm.

Argument	Value
Sparrows	10
Number of iterations	100
Training rounds	200
Batch size	60
Sequence input layer data dimension	240,2
Network parameters of the first layer 2-D convolution layer	5×1, 47, [2]
Network parameters of the second layer 2-D convolution layer	6×1, 63, [2]
Number of neurons in the LSTM layer	36
Output dimension of the fully connected layer	1

To address the issue of model overfitting and simulation data errors, this paper collects real-time data using collection cards to obtain rich datasets that enhance the model's generalization ability. Additionally, dropout regularization method is employed to prevent overfitting, and the model's complexity is constrained by adding a penalty term to the objective function. During the training process, the model is constantly modified and improved using the loss function [22]. The optimized parameters obtained through the above steps construct the automatic diagnostic model for short circuit faults in power meters based on the improved CNN-LSTM. The specific process of the model is shown in Figure 7.

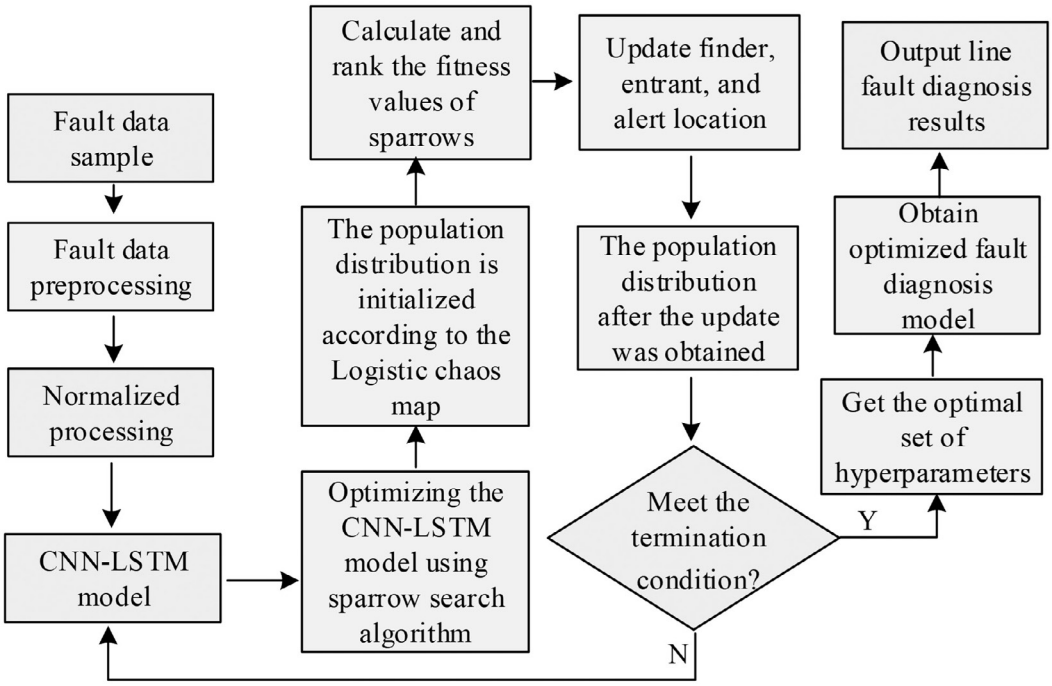


Figure 7. Flow of power measuring instrument short-circuit fault automatic diagnosis model.

4. Result and Discussion

4.1. Improvement Effect of Fault Diagnosis Model

DD238 and DB2 are single-phase and two-phase meter equipment, respectively, while DS864 and DT862 correspond to three-phase three-wire meter and three-phase four-wire meter. The experiment involved testing 600 samples of four types of power measuring instruments from the test set, with each type having 150 experimental samples. In the initial state, all meter samples were free of faults, and some samples were artificially set and damaged to introduce short circuit faults. The study used the sparrow search algorithm improved by the Logistic mapping for parameter optimization of the model. To verify the optimization effect, the training conditions of the model were compared and analyzed, including LSTM, CNN, and the unimproved CNN-LSTM. The specific comparison is shown in Figure 8.

According to Figure 8, from the change in the loss function, it can be observed that in the initial training stage, the generated samples do not possess the characteristic information contained in real samples [23]. The discriminator can easily distinguish between real and fake input samples, resulting in a large value of the loss function. After 834 iterations, the model has learned some features of the samples, and the loss function shows significant oscillations, indicating that the discriminative performance of the model is not yet stable [24]. After 1000 iterations, the loss function value of the improved CNN-LSTM approaches zero. Comparing the four models in the figure, the improved CNN-LSTM model has a significantly faster convergence speed. Compared to LSTM, CNN, and the unimproved CNN-LSTM, the iterations are reduced by 389, 408, and 218, respectively. This indicates that the sparrow search algorithm can effectively optimize the model and improve its convergence. To further validate the training effect, the study compares the training time, loss function value, training accuracy, and testing accuracy of the four models. The comparison is presented in Table 2.

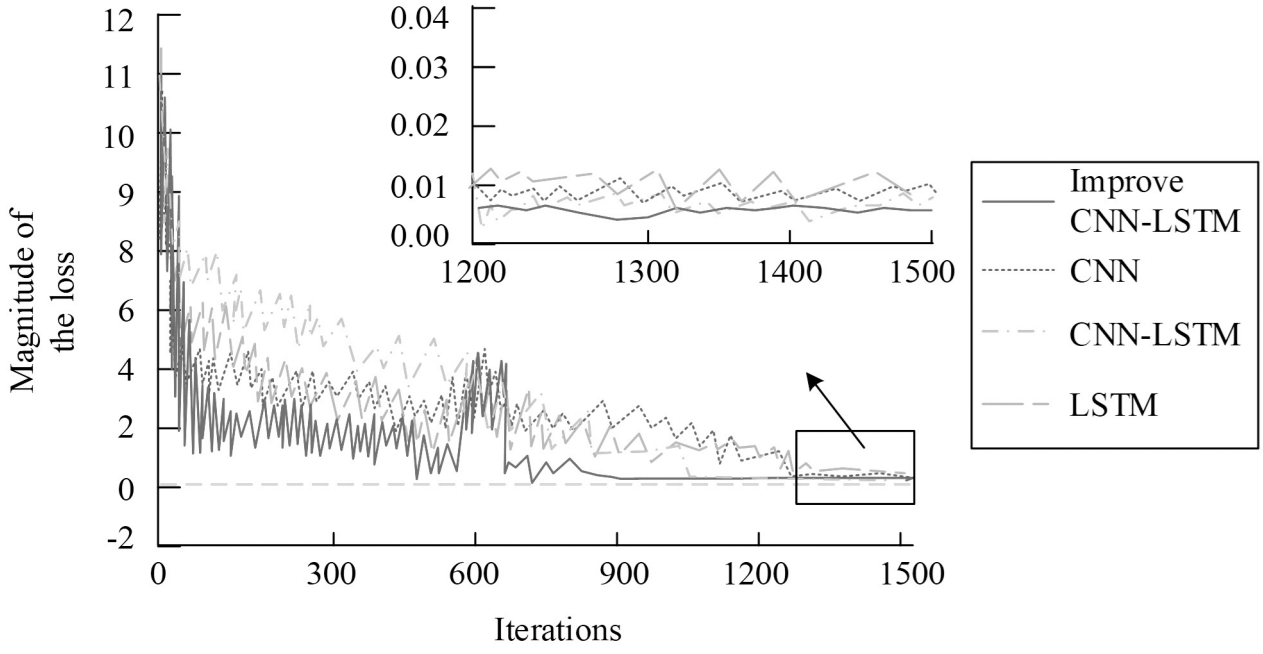


Figure 8. Loss function curves of different models.

Table 2. Training results comparing.

Project	Training time (s)	Loss function value	Training accuracy (%)	Test accuracy (%)
Improved CNN-LSTM	253.91	0.148	96.36	96.66
CNN-LSTM	384.26	0.281	92.14	91.77
CNN	436.02	0.346	89.77	89.44
LSTM	420.77	0.356	89.36	89.28

From Table 2, the improved CNN-LSTM mean training time is 253.91s. The average running time of CNN-LSTM, CNN, and LSTM is 384.26s, 436.02s, and 420.77s, respectively. Thus, the improved CNN-LSTM has the shortest training time. The average training accuracy of the improved CNN-LSTM reaches 96.36%, which is 7.00%, 6.59%, and 4.59% higher than that of LSTM, CNN, and the unimproved CNN-LSTM, respectively. The average testing accuracy of the improved CNN-LSTM reaches 96.66%, which is 7.38%, 7.22%, and 4.89% higher than that of LSTM, CNN, and the unimproved CNN-LSTM. The improved model has achieved gain in both diagnostic accuracy and efficiency. This is because the sparrow optimization algorithm optimizes the initial parameters of the model, allowing the model to achieve higher accuracy.

4.2. Model Feature Extraction Effect

The study uses wavelet packet energy spectrum and CNN for feature extraction of short circuit fault samples. To verify the effectiveness of feature signal extraction, the study uses the model to extract features from power meter signal data and compares them with the actual values [25]. The comparison results are shown in Figure 9.

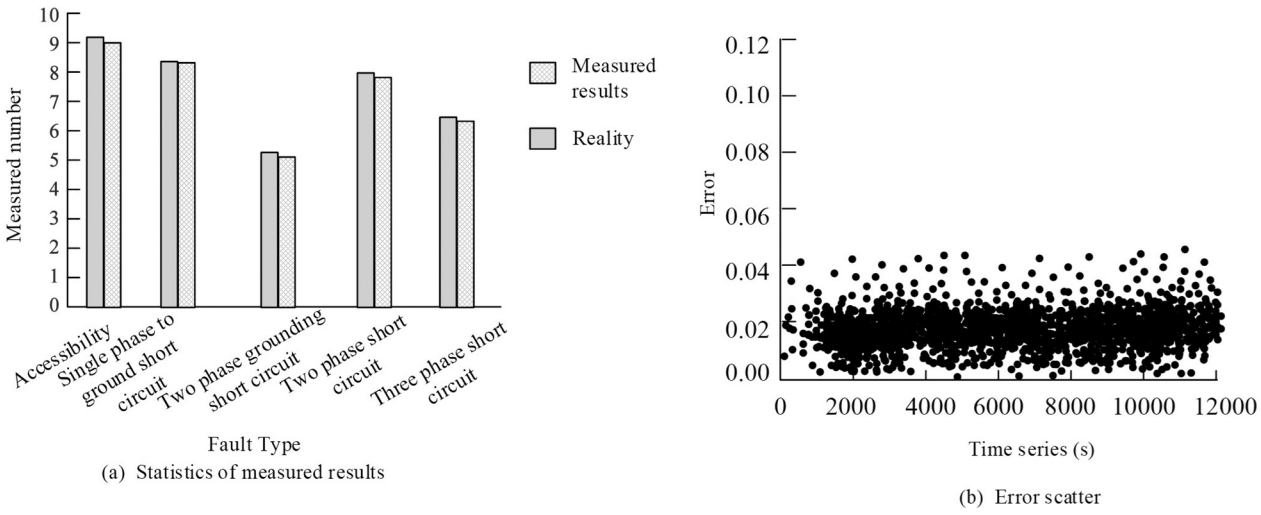


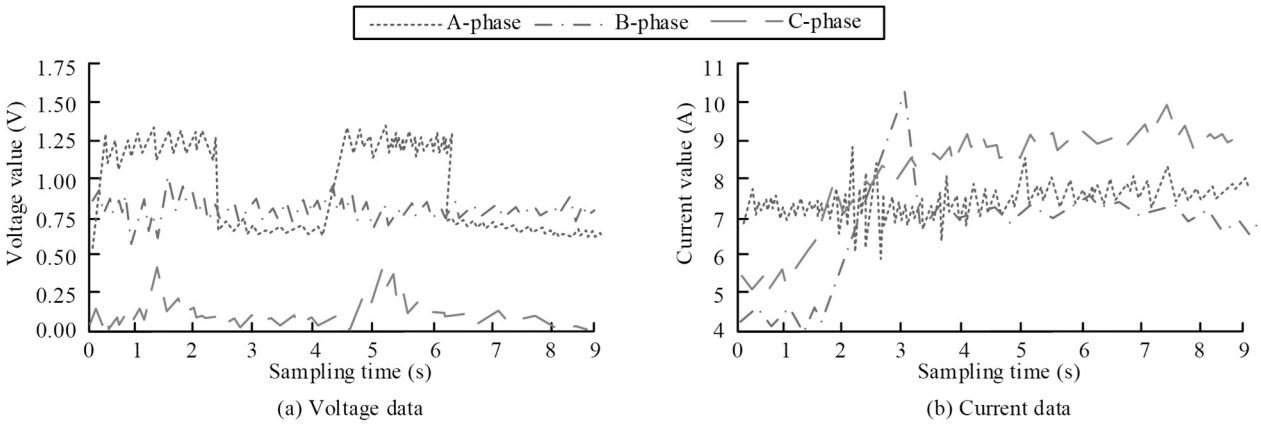
Figure 9. Comparison of model extraction features with measured values.

Utilizing wavelet energy spectrum distribution to extract signal features resulted in error values remaining in the range of 0.00 to 0.04 in comparison to the actual values, as shown in Figure 9. The majority of error values were concentrated around the value 0.02. This indicates that the feature extraction method used in the study can ensure the accuracy and reliability of the obtained power meter signal data, providing a more reliable basis for subsequent short circuit fault diagnosis [26]. The real-time operation data collection results of the research object 001, obtained through feature extraction, are shown in Figure 10.

From Figure 10, the operation data of all research objects can be obtained, which serves as the initial data for fault diagnosis, and ultimately leads to the final fault diagnosis results.

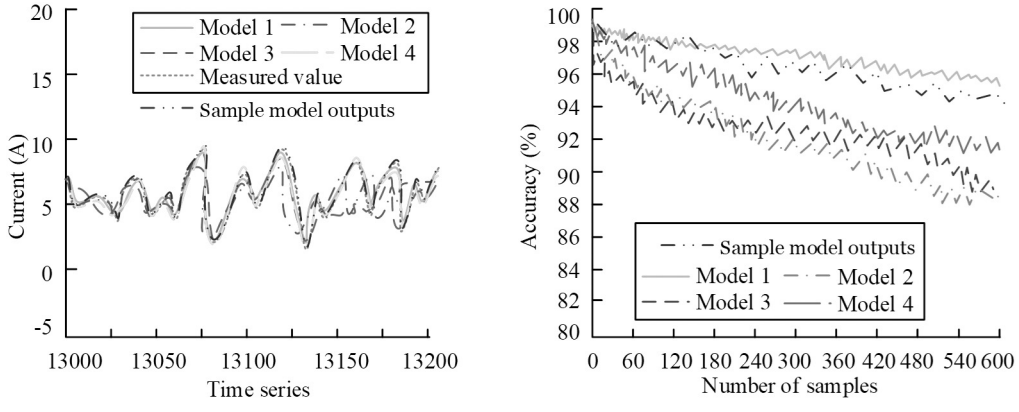
4.3. Performance Analysis of Fault Automation Diagnosis Model

To verify the diagnostic performance of the designed short circuit fault diagnosis model (Model 1) in the study, it is compared with three commonly used short circuit fault diagnosis models. The compared models include a SC fault diagnosis model based on fast Fourier transform and LSTM (Model 2), a SC fault diagnosis model based on coefficient deep forest (Model 3), and a SC fault diagnosis model based on temporal convolutional residual network (Model 4). Under different sample quantities, the four models are used to diagnose short circuit faults in the samples, and the diagnostic results for each type of short circuit fault are shown in Figure 11.



(a) Voltage data (b) Current data

Figure 10. Collection result of real-time operation data of 001 sample.



(a) Current measurement fitting results (b) Diagnostic accuracy

Figure 11. Comparison of short-circuit diagnosis of samples by the model.

Figure 11 shows that as the number of samples increases, the diagnostic accuracy of each model decreases to varying degrees. This is due to the gradual increase in sample size, the increase in model running load, and fluctuations in diagnostic performance. Model 1 has an average diagnostic accuracy of 96.46%, which is higher than the average diagnostic accuracy of the other three models by 13.24%, 10.89%, and 8.15%. Furthermore, Model 1 exhibited the smallest decrease in accuracy when the sample size increased from 10 to 600, with a decrease of only 2.14%. In contrast, the other three models all showed a decrease of over 5%. The change curve of Model 1 is consistent with the output of the sample model, as shown in Figure 11 (a), where the diagnostic result curve of Model 1 has a fitting degree of 0.92 with the output of the sample model, and the trend is consistent with the actual detection results. Based on the above content, Model 1 has the best diagnostic performance and good stability for power meter diagnosis. To further validate the performance of the model, the study applies the four models to diagnose different types of short circuit faults under the same number of samples. The obtained diagnostic results are shown in Table 3 [27–28].

From Table 3, Model 1 running time is 115.67s, 102.19s, and 88.64s less than that of Model 2, Model 3, and Model 4, respectively. The average diagnostic accuracy of Model 1 is 97.03%, which is 9.24%, 7.10%, and 4.25% higher than of the other models. The designed short circuit fault diagnosis model in the study can achieve real-time and high-precision fault diagnosis for power meters, ensuring quality provisioning of power supply services [29].

To assess the model performance comprehensively, the study rigorously benchmarks public datasets using an objective approach. The PEDL dataset is selected as the test data, and the same dataset is used to test each model to ensure consistency of test conditions. Objective metrics such as precision, recall rate, and F1 score are introduced before performing statistical significance tests. Multiple models are used concurrently for fault diagnosis and localization, and their test results are compared. The model is then optimized and improved based on the benchmark test results. The specific results of the benchmark tests are presented in Table 4.

Table 4 of the publicly available data set PEDL shows that Model 1 outperforms other models in terms of precision, recall rate, and F1 score, with statistically significant differences. This indicates that the short-circuit fault diagnosis model proposed in this paper has a superior diagnostic effect on the PEDL data set. However, other models still perform relatively well, albeit with a certain gap compared to Model 1. The running time of the models is analyzed to evaluate their performance in practical applications. The tests show that Model 1's running time decreases significantly compared to other models. These results demonstrate that the short-circuit fault diagnosis model developed by our research not only has high diagnostic accuracy but also has outstanding real-time performance and scalability, making it suitable for practical applications.

Table 3. Comparison of diagnostic results of the four models.

Model	Detection time (s)	Diagnostic accuracy (%)					
		Barrier free	Single-phase earth short circuit	Two-phase ground short circuit	Two-phase short circuit	Three-phase short circuit	Total
Model 1	226.78	96.88	97.08	98.99	96.44	95.78	97.03
Model 2	342.45	86.48	88.49	88.47	87.63	87.86	87.79
Model 3	328.97	90.45	89.46	89.77	89.88	90.10	89.93
Model 4	315.42	92.48	92.85	93.01	92.77	92.78	92.78

Table 4. Benchmarking results for the four models.

Project	Model 1	Model 2	Model 3	Model 4	*P*
Precision (%)	95.84	80.44	85.15	89.47	0.002
Recall (%)	94.88	80.87	85.67	90.17	0.001
F1	0.95	0.81	0.84	0.90	0.002
Detection time (s)	128.41	241.41	228.21	200.85	0.003

5. Conclusion

Fault diagnosis is a crucial method in power systems that power enterprises have recently prioritized. To improve the accuracy of short-circuit fault diagnosis in power measuring instruments, a CNN-based short-circuit fault diagnosis model has been developed. By utilizing CNNs and wavelet packet energy spectrum, significant signal features were identified. The Sparrow search algorithm and LSTM were used to optimize the model, resulting in automatic determination of fault type based on operational characteristic data.

The experimental results indicated that the iterations of CNN-LSTM optimized by the Sparrow search algorithm were reduced by 389, 408, and 218 times, respectively, compared to LSTM, CNN, and unimproved CNN-LSTM. This indicated that the sparrow search algorithm can significantly improve the model and significantly improve the convergence performance of the model. The CNN-LSTM model showed an average test accuracy of 96.66%, which was higher than the accuracy of LSTM, CNN, and unimproved CNN-LSTM by 7.38%, 7.22%, and 4.89%, respectively. Therefore, the enhanced model's diagnostic accuracy is significantly higher.

To conduct a comprehensive evaluation of the model, this study compared it with existing improved models. The results showed that Model 1 had a significantly shorter running time than Model 2, Model 3, and Model 4, with differences of 115.67s, 102.19s, and 88.64s, respectively. Additionally, Model 1 had a higher diagnostic accuracy, averaging at 97.03%, which represents an improvement of 9.24%, 7.10%, and 4.25% compared to the other three models. Model 1 displayed greater accuracy, recall rate, and F1 score compared to its counterparts on the PEDL public dataset. The observed disparities were statistically significant.

This study showed that the short-circuit fault diagnosis model designed in this paper performs better and has a better diagnostic effect on the public dataset. The model proposed in this paper efficiently and accurately diagnoses short circuit faults in power measuring instruments. This provides a basis for subsequent maintenance of the power system. Compared with traditional fault diagnosis methods, this model has higher accuracy and real-time performance, which can better meet the needs of modern power systems for fault diagnosis.

The study utilized the sparrow search algorithm to optimize short-circuit fault diagnosis models, resulting in improved convergence performance and diagnostic accuracy through continuous iteration and optimization. The study combines CNN with LSTM to extract fault features from power meters, and then further processes and analyzes these features by LSTM, finally achieving high-precision diagnosis of short-circuit faults. The model designed in this study has certain advantages in accuracy and runtime compared to other advanced models.

Deep learning-based fault diagnosis networks are diagnostic models trained on existing samples. The model presented in this paper has only been focused on short-circuit faults. However, in practical applications, there are still various types of faults, such as open circuit, overload, and others. Therefore, future research can further expand the model to achieve accurate diagnosis of more types of faults. Although the presented model has achieved some results, it still requires continuous optimization and improvement in practical applications. Future research can enhance the

model's performance by increasing the sample size, improving the network structure, and using other methods. Additionally, this model can be applied to other types of power equipment fault diagnosis to expand its application scope.

References

[1] M. T. Adkhar and H. Afianti, "Monitoring and Control System of Parallel Loads Electricity Consumption Based on IOT", *Journal of Electrical Engineering and Computer Sciences*, vol. 8, no. 1, pp. 63–70, 2023. http://dx.doi.org/10.54732/jeecs.v8i1.8

[2] M. M. Gomaa *et al.*, "Image Forgery Detection, Support Vector Machine (SVM), Deep Learning (DL), Convolutional Neural Networks (CNN), K-nearest Neighbor (KNN)", *Journal of System and Management Sciences*, vol. 12, no. 6, pp. 454–467, 2022. http://dx.doi.org/10.33168/JSMS.2022.0627

[3] R. Yang *et al.*, "On-board Diagnosis of Soft Short Circuit Fault in Lithium-ion Battery Packs for Electric Vehicles Using an Extended Kalman Filter", *CSEE Journal of Power and Energy Systems*, vol. 8, no. 1, pp. 258–270, 2020. http://dx.doi.org/10.17775/CSEEJPES.2020.03260

[4] W. Huang *et al.*, "Interturn Short-Circuit Fault Diagnosis of Interior Permanent Magnet Synchronous Motor for Electric Vehicle Based on Search Coil", *IEEE Transactions on Power Electronics*, vol. 38, no. 2, pp. 2506–2515, 2023. http://dx.doi.org/10.1109/TPEL.2022.3213512

[5] J. W. Choi and E. R. Jeong, "A Multi-output Convolutional Neural Network-based Distance and Velocity Estimation Technique", *Journal of Logistics, Informatics and Service Science*, vol. 9, no. 1, pp. 4555–4570, 2022. http://dx.doi.org/10.14704/WEB/V19I1/WEB19302

[6] A. Alloui *et al.*, "Online Interturn Short-Circuit Fault Diagnosis in Induction Motors Operating Under Unbalanced Supply Voltage and Load Variations, Using the STLSP Technique", *IEEE Transactions on Industrial Electronics*, vol. 70, no. 3, pp. 3080–3089, 2022. http://dx.doi.org/10.1109/TIE.2022.3172751

[7] A. M. Khelif, A. Bendiabdellah and B. D. E. Cherif, "Short-circuit Fault Diagnosis of the DC-Link Capacitor and Its Impact on an Electrical Drive System", *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 10, no. 3, pp. 2807–2814, 2020. http://dx.doi.org/10.11591/ijece.v10i3.pp2807-2814

[8] G. Wang *et al.*, "Voltage Correlation-Based Principal Component Analysis Method for Short Circuit Fault Diagnosis of Series Battery Pack", *IEEE Transactions on Industrial Electronics*, vol. 70, no. 9, pp. 9025–9034, 2022. http://dx.doi.org/10.1109/TIE.2022.3210588

[9] P. Zhang *et al.*, "A Novel Fault Diagnosis Technique of Interturn Short-circuit Fault for SRM in Current Chopper Mode", *IEEE Transactions on Industrial Electronics*, vol. 69, no. 3, pp. 3037–3046, 2021. http://dx.doi.org/10.1109/TIE.2021.3066917

[10] K. Zhang *et al.*, "An Early Soft Internal Short-Circuit Fault Diagnosis Method for Lithium-Ion Battery Packs in Electric Vehicles", *IEEE/ASME Transactions on Mechatronics*, vol. 28, no. 2, pp. 644–655, 2023. http://dx.doi.org/10.1109/TMECH.2023.3234770

[11] S. N. I. Zulkeefli and N. Hashim, "Comparison of CNN-based Algorithms for Halal Logo Recognition", *Journal of System and Management Sciences*, vol. 12, no. 5, pp. 155–168, 2022. http://dx.doi.org/10.33168/jsms.2022.0510

[12] Y. Xu and T. T. Qiu, "Human Activity Recognition and Embedded Application Based on Convolutional Neural Network", *Journal of Artificial Intelligence and Technology*, vol. 1, no. 1, pp. 51–60, 2020. http://dx.doi.org/10.37965/jait.2020.0051

[13] M. Wieczorek *et al.*, "Lightweight Convolutional Neural Network Model for Human Face Detection in Risk Situations", *IEEE Transactions on Industrial Informatics*, vol. 18, no. 7, pp. 4820–4829, 2021. http://dx.doi.org/10.1109/TII.2021.3129629

[14] D. Paraskevopoulos *et al.*, "Hybrid Wavelet–CNN Fault Diagnosis Method for Ships' Power Systems", *Signals*, vol. 4, no. 1, pp. 150–166, 2023. http://dx.doi.org/10.3390/signals4010008

[15] X. Pan *et al.*, "Deepopf: A Deep Neural Network Approach for Security-constrained DC Optimal Power Flow", *IEEE Transactions on Power Systems*, vol. 36, no. 3, pp. 1725–1735, 2020. http://dx.doi.org/10.1109/TPWRS.2020.3026379

[16] J. Jiang *et al.*, "Fault Diagnosis of the Bushing Infrared Images Based on Mask R-CNN and Improved PCNN Joint Algorithm", *High Voltage*, vol. 6, no. 1, pp. 116–124, 2020. http://dx.doi.org/10.1049/hve.2019.0249

[17] V. Nurmanova *et al.*, "Confidence Level Estimation for Advanced Decision-making in Transformer Short-circuit Fault Diagnosis", *IEEE Transactions on Industry Applications*, vol. 58, no. 1, pp. 233–241, 2021. http://dx.doi.org/10.1109/TIA.2021.3118661

[18] M. Ma *et al.*, "Fault Diagnosis Method Based on Multi-source Information Fusion for Weak Interturn Short Circuit in Synchronous Condensers", *IET Electric Power Applications*, vol. 15, no. 9, pp. 1245–1260, 2021. http://dx.doi.org/10.1049/elp2.12094

[19] R. A. C. Yohan *et al.*, "Identify Faults in Road Structure Zones with Deep Learning", *Journal of System and Management Sciences*, vol. 13, no. 1, pp. 63–84, 2023. http://dx.doi.org/10.33168/JSMS.2023.0104

[20] M. Tripathi, "Analysis of Convolutional Neural Network Based Image Classification Techniques", *Journal of Innovative Image Processing (JIIP)*, vol. 3, no. 2, pp. 100–117, 2021. http://dx.doi.org/10.36548/jiip.2021.2.003

[21] V. R. Allugunti, "A Machine Learning Model for Skin Disease Classification Using Convolution Neural Network", *International Journal of Computing, Programming and Database Management*, vol. 3, no. 1, pp. 141–147, 2021. http://dx.doi.org/10.33545/27076636.2022.v3.i1b.53

[22] S. Madhavi and S. P. Hong, "Anomaly Detection Using Deep Neural Network Quantum Encoder", *Journal of Logistics, Informatics and Service Science*, vol. 9, no. 2, pp. 118–130, 2022. http://dx.doi.org/10.33168/LISS.2022.0207

[23] W. Huang *et al.*, "New Recognition Method of Lightning Interference in Long Distance VSC-HVDC Transmission System", in *Proc. of the 2018 China International Conference on Electricity Distribution (CICED)*, 2018, pp. 1818–1823. http://dx.doi.org/10.1109/CICED.2018.8592017

[24] Y. Kim and W. Soh, "A Study on Character Recognition of Korean Vehicle License Plates Based on Deep Learning", *Journal of System and Management Sciences*, vol. 11, no. 1, pp. 69–82, 2021. http://dx.doi.org/10.33168/JSMS.2021.0105

[25] J. W. Choi and E. R. Jeong, "A Multi-output Convolutional Neural Network-based Distance and Velocity Estimation Technique", *Journal of Logistics, Informatics and Service Science*, vol. 9, no. 1, pp. 11–25, 2022. http://dx.doi.org/10.33168/LISS.2022.0202

[26] V. R. Allugunti, "A Machine Learning Model for Skin Disease Classification Using Convolution Neural Network", *International Journal of Computing, Programming and Database Management*, vol. 3, no. 1, pp. 141–147, 2021. http://dx.doi.org/10.33545/27076636.2022.v3.i1b.53

[27] S. Madhavi and S. P. Hong, "Anomaly Detection Using Deep Neural Network Quantum Encoder", *Journal of Logistics, Informatics and Service Science*, vol. 9, no. 2, pp. 118–130, 2022. http://dx.doi.org/10.33168/LISS.2022.0207

[28] J. Ma *et al.*, "A New Pilot Directional Protection Method for HVDC Line Based on the Polarity of Inductive Energy", *IET Generation,-Transmission& Distribution*, vol. 15, no. 22, pp. 3215–3228, 2021. http://dx.doi.org/10.1049/gtd2.12253

[29] I. H. Sarker, "Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions", *SN Computer Science*, vol. 2, no. 420, 2021. http://dx.doi.org/10.1007/s42979-021-00815-1

Received: October 2023
Revised: December 2023
Accepted: December 2023

Contact address:
Xiuni Xu
College of New Energy
Longdong University
Qingyang
Gansu
China
Tianjin University of Technology and Education
Tianjin
China
e-mail: xuxiuyixxy@163.com

Xiuni Xu received her master's degree from Tianjin University of Technology and Education in June 2023. At present, she is working as a lecturer in the College of New Energy, Longdong University. Her research is mainly focused on intelligent measurement and control systems, traffic information engineering and control.