

## Research Article

# Convolutional Neural Network for Voltage Sag Source Azimuth Recognition in Electrical Internet of Things

Ding Kai <sup>1</sup>, Li Wei,<sup>1</sup> Sun Jianfeng,<sup>2</sup> Xiao Xianyong,<sup>2</sup> and Wang Ying<sup>2</sup>

<sup>1</sup>Hubei Electric Power Research Institute State Grid, 430077 Wuhan, China

<sup>2</sup>College of Electrical Engineering, Sichuan University, 610065 Chengdu, China

Correspondence should be addressed to Ding Kai; shelliding@163.com

Received 23 October 2020; Revised 3 December 2020; Accepted 20 February 2021; Published 22 March 2021

Academic Editor: Shaohua Wan

Copyright © 2021 Ding Kai et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Recognition and analytics at the edge enable utility companies to predict and prevent problems in real time. Clearing the voltage sag disturbance source by the positioning method is the most effective way to solve and improve the voltage sag. However, for different grid structures and fault types, the existing methods usually achieve a sag source location based on the single feature of monitoring data extraction. However, due to the effectiveness and applicability of the existing method features, this paper proposes a multidimensional feature of the voltage sag source positioning method of the matrix. Based on the analysis of the characteristics of the voltage sag event caused by the fault, this paper proposes a multidimensional feature matrix for voltage sag source location, based on the convolutional neural network to establish the mapping relationship between the feature matrix and the voltage sag position, thus achieving multiple points based on multiple points. The voltage sag source orientation is identified by the monitoring data. Finally, the voltage sag event caused by the short-circuit fault is simulated in the IEEE14 node model, and the effectiveness of the proposed method is verified by simulation data. The simulation results show that the proposed method has higher accuracy than the traditional method, and the method can be applied to different grid structures and different types of faults.

## 1. Introduction

Electric Internet of Things (EIIoT) realizes the interconnection of everything and human-computer interaction in all aspects of the power system by fully applying mobile Internet technology, artificial intelligence technology, and advanced communication technology. It provides intelligent services characterized by a comprehensive perception of the state, efficient processing of information, and convenient and flexible application. In order to drive reliability and operational efficiency, utility companies now need to leverage advanced distribution analysis. Delivering analytics at the edge means that the energy industry can use data in a far more effective and efficient way. Detection at the edge enables utility companies to predict and prevent problems in real time, to cost-effectively deploy resources and personnel, and to increase overall grid optimization, security, and reliability. A voltage sag is a power quality disturbance event where supply voltage quickly recovers after dropping [1]. The root

cause of the voltage sag is the voltage sag of a common connecting point which is caused by the increase in the partial voltage of power impedance due to the increase in current. The causes of the voltage sag can be divided into short-circuit fault, transformer magnetizing inrush current, inductive motor start, and other large-capacity load operations [2–4]. The positioning of the voltage sag can not only be used to divide the temporary responsibility of both power supply and power consumption but also help to shorten the time spent by a power supply company on fault elimination [5].

At present, the positioning methods of the current voltage sag can be roughly divided into positioning methods based on a single monitoring point [6–9] and positioning methods based on multiple monitoring points [10, 11]. Literature [6] determines whether the voltage sag source is in the upstream or downstream of the power quality detection device according to the signs of the disturbed active power (i.e., the difference between the active power before and after

the occurrence of the voltage sag) and the disturbed active power energy (the integral of the disturbed active power on a time scale). Literature [7] uses the fundamental voltage and current recorded by the power quality monitoring device to calculate equivalent impedance and determines whether the voltage sag is located in the upstream or downstream of the power quality monitoring device according to the signs of the real part of impedance. Literature [8] judges that the voltage sag source is located in the upstream or downstream of the monitoring point according to the slope of the line segment connected by two points before and after the fault. Literature [9] determines whether the voltage sag source is located in the upstream or downstream of the monitoring point according to the signs of the real part of current during the voltage sag. Literature [10] puts forward the sensitivity index of the fault current to the voltage sag, determines the priority of nodes according to the sensitivity index, and then determines the priority of the installation location of the power quality monitor according to this index. Then, the position of the voltage sag source is determined according to the current variation of each monitoring branch before and after the voltage sag event as well as the priority of the nodes. Literature [11] searches the voltage sag path by using the feeder sending end and determines that the branch line with the lowest voltage is the line where the voltage sag occurs. Then, a quadratic function is used to fit the voltage into the function of distance, which can not only realize the estimation of node voltage sag but also achieve the location of the voltage sag source.

Traditional locating methods of the voltage sag source have some limitations in practical application. The locating methods based on a single monitoring point are only suitable for a radiation network. Besides, the judgment results based on the disturbed active power and disturbed active energy may not match [2], and this method lacks theoretical derivation [12]. The real part method of equivalent impedance [7], the system trajectory slope method [8], and the real part current method [9] all assume that the circuit conditions will not change when there is a voltage sag caused by the fault, which is bound to affect the judgment result. Although the method proposed in literature [10] is applicable to radiation networks and ring networks, it is affected by the system grounding mode. The method proposed in literature [11] is only applicable to radial networks. To solve the problems such as limited application scope and inaccurate positioning of the existing methods, literature [3, 13, 14] integrates the system slope trajectory and other traditional positioning methods as comprehensive criteria and uses the BP Neural Network (BPNN) and Support Vector Machine (SVM) to establish the mapping relationship between the criteria and positioning results, respectively. Literature [15, 16] obtains the first peak of the disturbed power after the occurrence of the voltage sag as the feature vector, and then, the SVM and Radial Basis Function (RBF) Neural Network are used to conduct classification and then achieve fault location, respectively. The voltage sag source location method based on classification considers the voltage sag source location a binary classification problem and determines whether the voltage sag source is located in the upstream or downstream of a single

monitoring point. Such methods are prone to have conflicting features that affect the classification results, especially when they are used for voltage sag positioning in a loop network.

As the power grid fault is the main cause of the voltage sag [17, 18], in this paper, on the basis of the research of the existing methods, the main contributions are as follows:

- (1) A voltage sag source location method based on a multidimensional feature matrix is proposed to conduct the positioning of the voltage sag events caused by the power grid fault
- (2) Then, the characteristics of voltage sag events are analyzed, and the disturbance power, disturbance energy, real part of equivalent impedance, slope trajectory, and real part of the current of all monitoring nodes of the whole network are extracted to form the feature matrix
- (3) Later, the locating problem of the voltage sag source is transformed into a multiclassification problem. The features are extracted by the convolutional layer and sampling layer of the convolutional neural networks (CNN) to avoid the contradiction between the features. The mapping relationship between deep features and the location of the voltage sag source is established through the fully connected layer of CNN to realize the location of the voltage sag source
- (4) In the IEEE14 node model, the grid fault data is obtained by simulation, in order to verify the effectiveness of the proposed method. The simulation results show that the proposed method is more accurate than the traditional methods and more applicable to different network structures and different types of faults

The rest of the paper is organized as follows. We first discuss some related works in voltage sag source location methods, and then, we elaborate our proposed sag source location method based on convolutional neural network algorithm. Moreover, we present experiments to evaluate our method in Simulation Analysis. Finally, the conclusions of this paper are given in Conclusion.

## 2. Voltage Sag Source Location Methods

*2.1. Upstream and Downstream Positioning Methods.* Both positioning criteria [6–9] proposed according to the variation of electric quantity when the voltage sag occurs and the method that constitutes a comprehensive criterion for a classifier's classification and identification [13–16] judge whether the voltage sag source is located in the upstream or downstream of the power quality monitoring device. As shown in Figure 1, the reference direction of the power flow has been marked with arrows. As for the power quality monitoring device  $M$ , the lines  $L_1$  and  $L_4$  where the energy flows to  $M$  and the lines  $L_2$ ,  $L_3$ , and  $L_{10}$  where the energy does not flow through  $M$  are the upstream regions, while the lines  $L_5 \sim L_9$

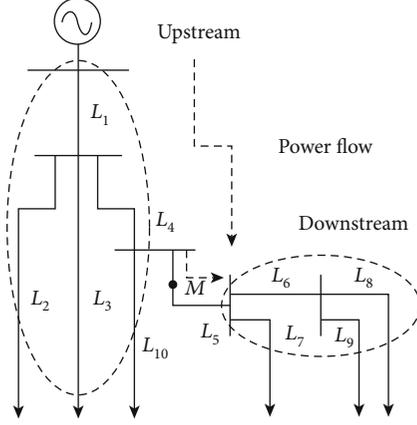


FIGURE 1: Voltage sag source location.

where the energy flows through  $M$  are the downstream regions. This method which defines upstream and downstream is only applicable to radial networks, but for more complex networks such as ring networks, it may be difficult for this method to define upstream and downstream.

**2.2. Multimonitoring Point Positioning Method.** The positioning method based on multimonitoring points needs to install a large number of monitoring devices in the power grid. At present, the positioning methods based on multimonitoring points mainly include the deviation method based on branch current [10] and the node voltage deviation index method [11].

The branch current deviation method uses the system coefficient matrix to calculate the current of the system when the bus  $k$  fails:

$$\begin{aligned} f(V_f^k) &= \sum_{i=1}^m \sum_{\substack{j=1 \\ j \neq i}}^m I_{ij}^k \\ &= \sum_{i=1}^m \sum_{\substack{j=1 \\ j \neq i}}^m \frac{(z_{ik} - z_{jk})(V_f^k / (z_{kk} + z_f^k))}{z_{b,ij}}, \end{aligned} \quad (1)$$

where  $z_{kk}$  is the equivalent resistance at the bus  $k$ ,  $z_f^k$  represents the fault resistance,  $z_{ik}$  is the transmission impedance between the buses  $i$  and  $k$ , and  $z_{b,ij}$  represents the line impedance between the buses  $i$  and  $j$ .

The sensitivity of the system current to the transient voltage is defined as

$$\frac{\partial f}{\partial V_f^k} = \sum_{i=1}^m \sum_{\substack{j=1 \\ j \neq i}}^m \frac{(z_{ik} - z_{jk})}{z_{b,ij}(z_{kk} + z_f^k)}. \quad (2)$$

The branch current deviation is

$$\text{BCD} = \frac{I_{\text{sag}} - I_{\text{pre}}}{I_{\text{pre}}}, \quad (3)$$

where  $I_{\text{pre}}$  and  $I_{\text{sag}}$  are defined as the current values before and after the failure, respectively.

The sensitivity of all nodes in the network is calculated, and the monitoring device is installed in the nodes with higher sensitivity. When a fault occurs, the current deviation of each branch of the bus equipped with the monitoring device is calculated, and the positioning of the voltage sag source can be realized by searching according to the current deviation. This method has obvious defects, and the fault current varies in different grounding modes, so it is difficult to guarantee the reliability of its application.

The node voltage deviation method calculates the voltages of all nodes in the whole network based on the system coefficient and monitoring data. The voltage deviation is defined as

$$V = \frac{V_{\text{pre}} - V_{\text{sag}}}{V_{\text{pre}}} \times 100\%. \quad (4)$$

The bus  $V_{\text{max}} = [V_1, V_2, V_3 \dots V_n]$  with the maximum voltage deviation is determined as the location where the voltage sag source occurs. This method does not take into account the voltage sag caused by different short-circuit faults, so it has some limitations.

### 3. Multidimensional Characteristic Matrix for Sag Source Location

Effective feature selection is the key to realize classification. Based on the existing methods, the disturbance power and energy, the real part of the equivalent impedance, the slope of the system trajectory, and the real part of the current are extracted from each node.

**3.1. Disturbance Power and Disturbance Energy.** DP (disturbance power) is defined as the difference between the instantaneous power and the steady-state operating power, while DE (disturbance energy) is defined as the integral of the disturbance energy during the disturbance.

$$\text{DP} = p(t)_f - p(t)_s, \quad (5)$$

$$\text{DE} = \int_0^t \text{DP}(t) dt, \quad (6)$$

where  $p(t)_f$  and  $p(t)_s$  represent the transient power during disturbance and steady-state operation, respectively.

**3.2. Equivalent Impedance Real Part.** Based on the equivalent circuit shown in Figure 2, the equivalent impedance variation at the power quality monitoring device caused by the voltage sag is analyzed.

Assuming that there is a disturbance in the downstream of the power quality monitoring device, the system

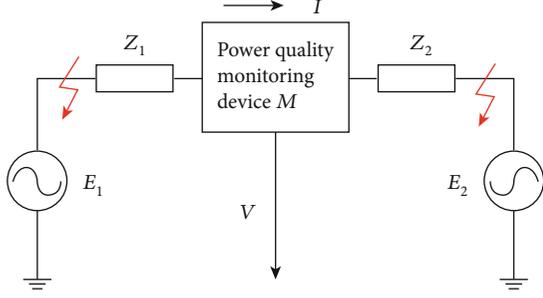


FIGURE 2: Equivalent circuit of the voltage sag.

parameters do not change during the disturbance, the voltage variation is  $\Delta V$ , and the current variation is  $\Delta I$ , then

$$V = E_1 - IZ_1, \quad (7)$$

$$V + \Delta V = E_1 - (I + \Delta I)Z_1. \quad (8)$$

The equivalent impedance can be calculated by Equations (7) and (8):

$$Z_2 = \frac{\Delta V}{\Delta I}. \quad (9)$$

Similarly, the equivalent impedance can be calculated when the disturbance occurs in the upstream:

$$Z_1 = -\left(\frac{\Delta V}{\Delta I}\right). \quad (10)$$

Therefore, the equivalent impedance can be defined as  $Z_e = \Delta V / \Delta I$ , and the voltage sag source can be localized according to the polarity.

### 3.3. System Slope Trajectory.

$$VI \cos \theta_2 = -I^2 \operatorname{Re}(Z_1) + E_1 I \cos \theta_1. \quad (11)$$

Equation (10) can be transformed into

$$V \cos \theta_2 = -I \operatorname{Re}(Z_1) + E_1 \cos \theta_1, \quad (12)$$

where  $\theta_2$  denotes the phase difference between  $V$  and  $I$  and  $\theta_1$  represents the phase difference between  $E$  and  $I$ .

When the disturbance occurs in the upstream of the power quality monitoring device,  $\cos \theta_2 < 0$ ; when the disturbance occurs in the downstream of the power quality monitoring device,  $\cos \theta_2 > 0$ . It can be seen from Equation (11) that, assuming that the operating parameters do not change when the disturbance occurs, the position of the voltage sag can be judged by the correlation between  $V \cos \theta_2$  and  $I$ . Therefore, a line is synthesized during the disturbance to locate the voltage sag source according to the slope of the line  $k$ .

**3.4. Real Part of the Current.** According to Equation (12), when the disturbance occurs in the upstream, the current

direction during the disturbance is opposite to that before the disturbance. When the disturbance occurs in the downstream, the current direction during the disturbance is the same as that before the disturbance. Therefore, the sag source can be localized according to the polarity.

**3.5. Characteristic Matrix.** The characteristic matrix of the whole network is as follows:

$$M_i = \begin{bmatrix} DP_1 & DP_2 & \cdots & DP_n \\ DE_1 & DE_2 & \cdots & DE_n \\ Ze_1 & Ze_2 & \cdots & Ze_n \\ k_1 & k_2 & \cdots & k_n \\ I \cos \theta_1 & I \cos \theta_2 & \cdots & I \cos \theta_n \end{bmatrix}, \quad (13)$$

where  $M_i (i \in \{a, b, c\})$  represents the characteristic matrix of phase  $i$ , and the three-dimensional characteristic matrix formed is used as the input of the three channels of CNN.

## 4. Sag Source Location Method Based on the Convolutional Neural Network (CNN)

The voltage sag location method proposed in this paper is used to classify and identify voltage sag events occurring on different lines. When the 3D characteristic matrix selected in Section 1 is used for classification and recognition, there may be a problem of mutual redundancy among features or failure to completely represent the voltage sag, which affects the accuracy of the final classification. As a bionic model, CNN's basic framework usually includes a feature extractor and classifier [19], and it is composed of a convolutional layer, sampling layer, and fully connected layer [20–22]. Its structure is shown in Figure 3. The convolutional layer and the sampling layer can not only effectively learn the original feature matrix of the input at a deeper level but also reduce the number of neurons and simplify the complexity of the network through weight sharing. The fully connected layer, as a classifier, inputs the features after convolution and sampling and outputs classification categories.

**4.1. Forward Pass.** In the forward pass process, the output of the  $l$ th convolutional layer is

$$x_j^l = f \left( \sum_{i \in M_j} x_i^{l-1} * k_{ij}^l + b_j \right), \quad (14)$$

where  $x_j^l$  denotes the  $j$ th output characteristic matrix of the  $l$ th layer,  $f$  denotes the activation function,  $M_j$  denotes the combination of the characteristic matrix output by the upper layer,  $k_{ij}^l$  denotes the convolution kernel connecting  $x_i^l$  and  $x_i^{l-1}$ , and  $b_j$  denotes the bias corresponded by  $x_j^l$ .

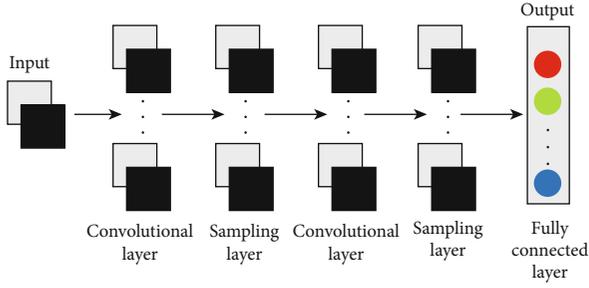


FIGURE 3: The structure of CNN.

The output of the  $l$ th convolutional layer is the input of the  $l$ th sampling layer. In this case, the output of the  $l$ th sampling layer is

$$x_j^l = \text{down}(x_j^{l-1}) + b_j^l, \quad (15)$$

where  $\text{down}()$  represents the lower sampling function.

The output obtained by convolution sampling is sorted into  $X = \{x_1, x_2, \dots, x_n\} \in \mathbb{R}^n$ , as the input of the fully connected layer, and the output of the  $l$ th fully connected layer is

$$X^l = f(W_{l-1}^l X^{l-1} + b^l), \quad (16)$$

where  $W_{l-1}^l$  denotes the weight connecting  $X^{l-1}$  and  $X^l$  and  $b^l$  is the bias of the  $l$ th layer.

**4.2. Backward Pass.** The training of CNN minimizes the error between the output data and the expected data by adjusting the weight and bias, and this process is known as reverse pass. As a supervised learning algorithm, it is necessary to define a loss function before conducting reverse pass. The commonly used error functions include the square error loss function and cross-entropy loss function. In this paper, the cross-entropy loss function is selected:

$$\text{Loss}(W, b, x, y_-) = -\frac{1}{n} \sum [y_- \cdot \ln y + (1 - y_-) \cdot \ln(1 - y)], \quad (17)$$

where  $n$  represents the number of samples,  $y_-$  represents the expected output, and  $y$  represents the actual output.

According to the loss function, the error of each neuron in the  $l$ th output layer can be calculated:

$$\delta_j^l = \frac{\partial L}{\partial X_j^l} f'(z_j^l), \quad (18)$$

where  $z_j^l$  is the value of the input activation function of the  $j$ th neuron in the  $l$ th layer. Equation (17) can be written as the matrix form as follows:

$$\delta^l = \nabla_X L \odot f'(Z^l), \quad (19)$$

where  $\odot$  represents the Hadamard product.

The error of the  $l$ th layer of the fully connected layer can be used to calculate the error of the  $l-1$ th layer of the fully connected layer:

$$\delta^{l-1} = \left( (W^l)^T \delta^l \right) \odot f'(Z^{l-1}). \quad (20)$$

Given the sampling layer error, the error of the previous layer can be calculated as follows:

$$\delta^{l-1} = \text{upsample}(\delta^l) \odot f'(Z^{l-1}), \quad (21)$$

where  $\text{upsample}()$  represents the upsampling function.

Given the convolution layer error, the error of the previous layer can be calculated as follows:

$$\delta^{l-1} = \delta^l * \text{rot180}(W^l) \odot f'(Z^{l-1}). \quad (22)$$

The backpropagation of the error can be used to calculate the gradient of the weight and bias:

$$\begin{aligned} \frac{\partial L}{\partial w_{jk}^l} &= Z_k^{l-1} \delta_j^l, \\ \frac{\partial L}{\partial b_j^l} &= \delta_j^l. \end{aligned} \quad (23)$$

The gradient descent method is adopted to adjust the weight and bias of the network:

$$Wb = Wb - \beta \frac{\partial L}{\partial Wb}, \quad (24)$$

where  $Wb$  denotes the weight or bias and  $\beta$  denotes the learning rate.

**4.3. Sag Source Positioning Method.** In this paper, a voltage sag source positioning method based on a multidimensional characteristic matrix is proposed to locate single voltage sag events caused by the grid fault. The positioning of the voltage sag source is considered a multiclassification problem, and the voltage sag events occurring on different lines are divided into different categories. The mapping relationship between voltage sag events and lines is established by CNN to realize the location of the voltage sag source. The flowchart of the method proposed in this paper is shown in Figure 4.

- (1) Record the voltage and current waveform of all monitoring points in the whole network when the voltage sag occurs. The IEEE14 node model is built in PSCAD/EMTDC to simulate voltage sag events caused by the short-circuit fault, and then, the voltage waveform of each node corresponded by each voltage sag event and the current waveform at both ends of each line are recorded
- (2) Based on the data recorded in step (1), the disturbance power, disturbance energy, real part of

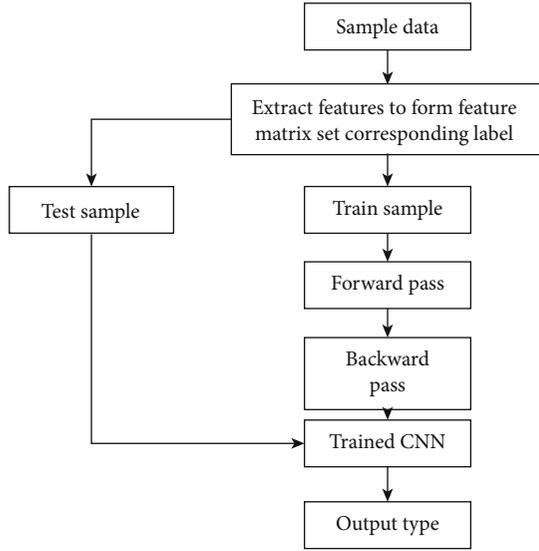


FIGURE 4: Flowchart of the method.

equivalent impedance, slope of system trajectory, and real part of the current at each end of each line are calculated to form a feature matrix, and then, the corresponding labels are set according to the line where the voltage sag occurs. The characteristic matrix and corresponding tags are used as a sample for CNN training. In this paper, the CNN trained for the test randomly selects a part of the obtained samples as the test samples

- (3) Take the training sample as the input of CNN and conduct forward pass to output the predicted tag. In order to make CNN output the result that matches the actual tag, reverse pass is required to adjust the network parameters. The trained CNN has established the mapping relationship between input and output
- (4) Based on the CNN completed by training, the positioning of the voltage sag source can be realized. Input the characteristic matrix of voltage sag events that need to be located, and the output is the label of the circuit that occurs. This paper tests the classification performance of CNN completed by training based on the test samples obtained in step (2).

## 5. Simulation Analysis

**5.1. Data Acquisition.** To verify the CNN-based voltage sag source location method proposed in this paper, a large amount of data of voltage sag events are needed. Firstly, the IEEE14 node model is built in PSCAD/EMTDC, as shown in Figure 5. Faults are set at each ten equal points of each line, including single line-to-ground fault (SLGF), phase-to-phase fault (PPF), double line-to-ground fault (DLGF), and three line-to-ground fault (TLGF). Ten uniform random numbers whose transition resistance is between 0 and 10  $\Omega$  are set at each fault point for simulation. Then, the sampling frequency of 5 kHz is used to record the voltage waveform of 14 nodes

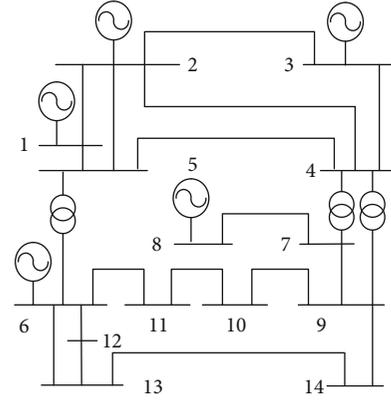


FIGURE 5: IEEE14 model.

TABLE 1: Distribution of samples.

Type	Sample size	Training sets	Testing sets
SLGF	1700	1360	340
PPF	1700	1360	340
DLGF	1700	1360	340
TLGF	1700	1360	340

and the current waveform at both ends of 17 lines during each simulation.

The recorded waveform data is input into MATLAB for processing. First of all, an FIR low-pass filter is designed to filter the collected data and retain the power frequency component. Secondly, the features mentioned in Section 2 at both ends of each line are extracted to form a feature matrix of  $5 \times 34 \times 3$ , and then, the corresponding label (label format is  $5 \times 34 \times 3$ ) is set. The 100 sample sets of each line are randomly divided into 80 training samples and 20 test samples, and the sample distribution is shown in Table 1.

**5.2. Model Evaluation Index.** The indexes which are used to evaluate dichotomy problems mainly include accuracy rate, recall rate, F1 value index, ROC curve, etc. [23–26], but they are no longer fully applicable to the multiple classification problems appearing in this paper. Therefore, the accuracy rate and kappa coefficient [27–30] are intended to be used to evaluate the models used in this paper.

The calculation formula of kappa is as follows:

$$K = \frac{P_0 - P_e}{1 - P_e}, \quad (25)$$

$$P_e = \frac{a_1 b_1 + a_2 b_2 + \dots + a_i b_i}{N \cdot N}, \quad (26)$$

where  $P_0$  represents the classification accuracy of samples,  $a_i$  represents the number of samples of category,  $b_i$  represents the number of samples predicted to be category  $i$ , and  $N$  represents the total number of samples.

**5.3. Construction of the CNN Model.** Based on TensorFlow, a deep learning framework developed by Google, CNN is built.

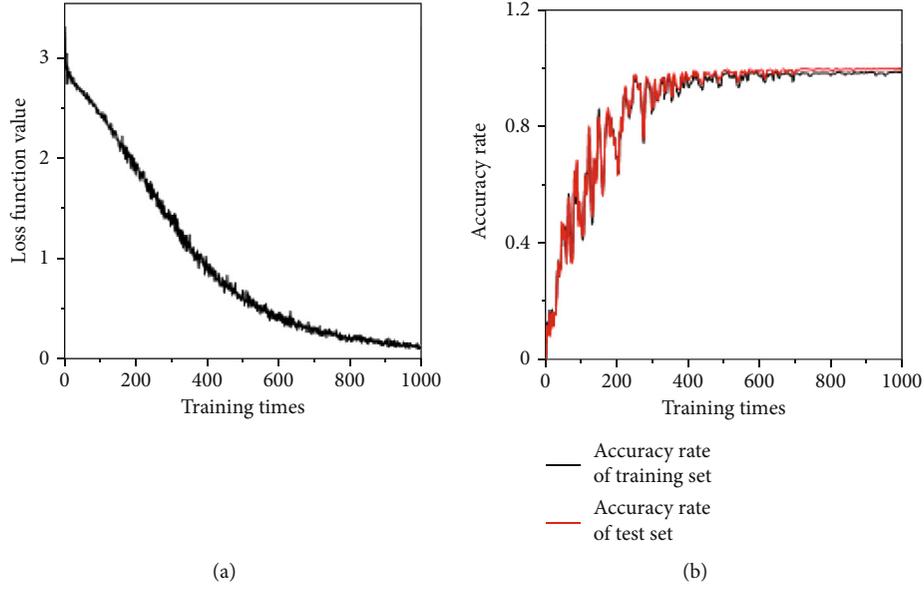


FIGURE 6: Training process: (a) loss function value curve and (b) accuracy rate curve.

L1_2	20.00	0.00	0.00	0.00	0.00
L2_3	0.00	19.00	0.00	1.00	0.00
L2_4	2.00	0.00	18.00	0.00	⋮
L6_13	0.00	0.00	0.00	20.00	0.00
...	0.00	0.00	...	0.00	20.00
	L1_2	L2_3	L2_4	L6_13	...

FIGURE 7: Confusion matrix of classification results.

The CNN consists of two convolutional sampling layers and three fully connected layers [31–36]. The first convolutional layer has 32 convolution kernels with a size of  $2 \times 5$ , the second convolutional layer has 64 convolution kernels with a size of  $1 \times 6$ , the convolution kernel of the sampling layer is  $2 \times 2$ , and the number of nodes in the middle layer of the fully connected layer is 1024. Among them, the activation function of the hidden layer is “RELU,” the activation function of the output layer is “SOFTMAX,” the convolution mode is “VALID,” and the sampling mode is “SAME.”

The CNN which has been built is trained in batches, and 80 training samples are randomly trained in each batch. The weight is set as a normally distributed random number with a mean value of 0 and variance of 0.1, the bias is set as a constant of 0.1, the learning rate is set as 0.0001, and the number of training is set as 1000. Adam-Optimizer is selected to

TABLE 2: Classification indicator.

Classification algorithm	Classification feature	Accuracy	Kappa
CNN	Characteristic matrix	99.12%	99.06%
	Disturbance power	97.06%	96.88%
	Equivalent impedance	43.53%	40.00%
	System slope	51.76%	48.75%
KNN	Real part of current	57.94%	55.31%
	Characteristic matrix	87.94%	87.19%
	Disturbance power	40.29%	36.56%
	Equivalent impedance	57.06%	54.38%
EL	System slope	41.18%	37.50%
	Real part of current	32.94%	28.75%

adaptively control the learning rate of each parameter in the network.

**5.4. Single-Phase Ground Fault Analysis.** The classification results of the voltage sag source caused by SLGF are analyzed, and only the classification results of other fault types are given. The change curve of the loss function value and the change curve of classification accuracy of the training set and test set are shown in Figure 6.

Based on the obfuscation matrix, the evaluation index of the model is calculated, each column of the obfuscation matrix represents the predicted category, and each row represents the actual category. In this paper, there are 17 types of SLGF samples, so the dimension of the confusion matrix is  $17 \times 17$ . To observe and draw the confusion matrix easily, as shown in Figure 7, the omitted parts are all correctly classified samples. Based on the confusion matrix, one sample of line L2\_3 (L2\_3 represents the line between node 2 and node

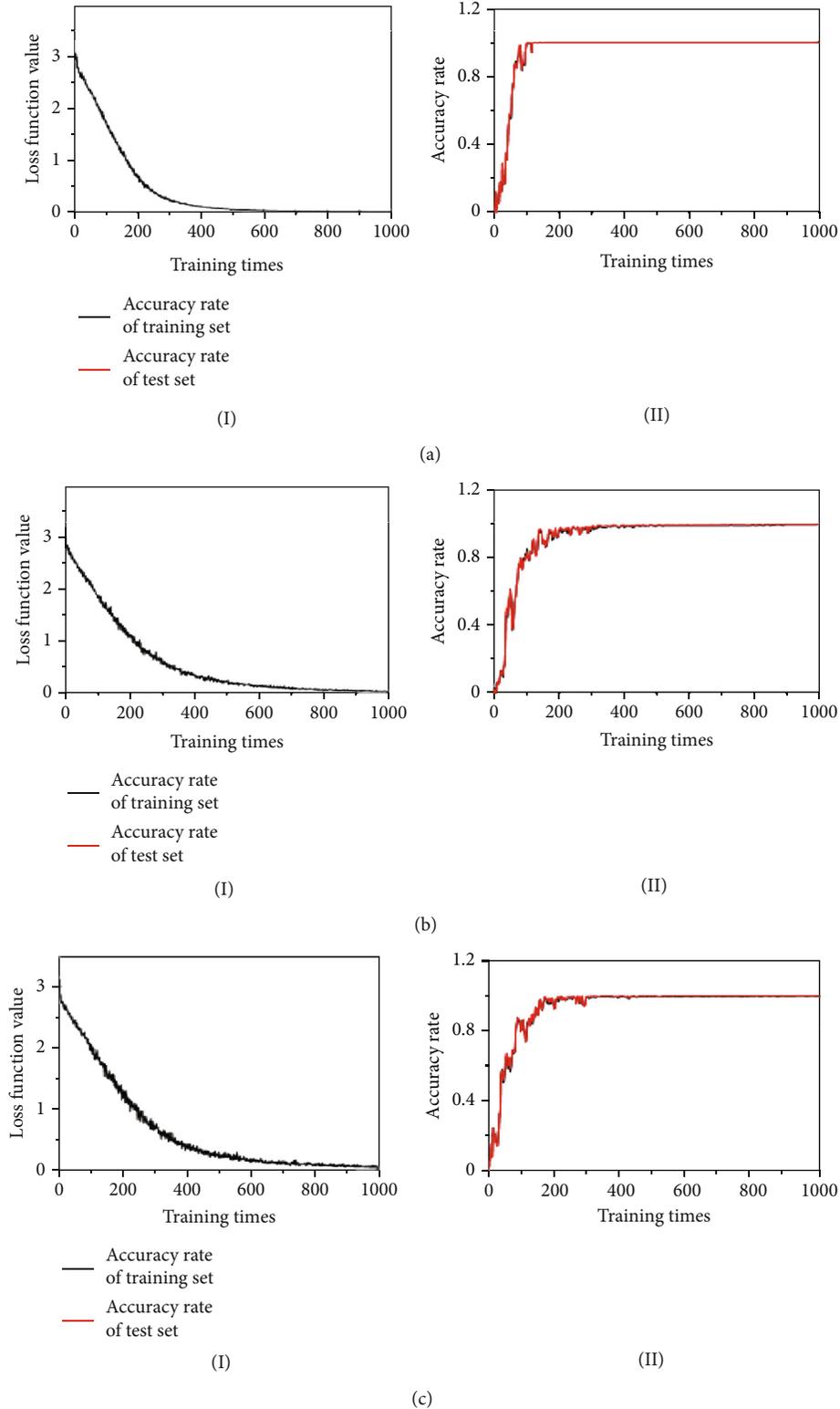


FIGURE 8: Training process: (a) phase-to-phase fault, (b) double line-to-ground fault, and (c) three line-to-ground fault.

3) is misclassified as line  $L6-13$  by the classifier and two samples of line  $L-4$  as line  $L1-2$ , and the samples of other lines can be correctly classified as the real category.

The confusion matrix as shown in Figure 7 can be drawn by using the  $K$ -nearest Neighbor (KNN) and Ensemble

Learning Algorithm (EL) (the confusion matrix of the classification results of the comparison classifier is omitted here due to space limitation). The respective classification indexes can be calculated according to the confounding matrices of different classifiers, as shown in Table 2.

TABLE 3: Classification indicator.

Fault type	Classification algorithm	Classification feature	Accuracy	Kappa
PPF	CNN	Multidimensional feature	100.00%	100.00%
		Multidimensional feature	83.24%	82.19%
		Disturbance power	88.53%	87.81%
	KNN	Equivalent impedance	76.18%	74.69%
		System slope	82.35%	81.25%
		Real part of current	82.35%	81.25%
	EL	Multidimensional feature	64.12%	61.88%
		Disturbance power	64.41%	62.19%
		Equivalent impedance	64.12%	61.88%
		System slope	64.71%	62.50%
DLGF	CNN	Multidimensional feature	100.00%	100.00%
		Multidimensional feature	90.59%	90.00%
		Disturbance power	93.24%	92.81%
	KNN	Equivalent impedance	82.35%	81.25%
		System slope	89.41%	88.75%
		Real part of current	81.18%	80.00%
	EL	Multidimensional feature	63.82%	61.56%
		Disturbance power	58.82%	56.25%
		Equivalent impedance	64.41%	62.19%
		System slope	64.12%	61.88%
TLGF	CNN	Multidimensional feature	100%	100%
		Multidimensional feature	97.65%	97.50%
		Disturbance power	79.12%	77.81%
	KNN	Equivalent impedance	80.00%	78.75%
		System slope	31.18%	90.63%
		Real part of current	76.76%	75.31%
	EL	Multidimensional feature	64.12%	61.88%
		Disturbance power	35.29%	31.25%
		Equivalent impedance	35.29%	31.25%
		System slope	47.06%	43.75%
		Real part of current	64.12%	61.88%

The classification accuracy of CNN is 99.12% and the kappa coefficient is 99.06%, which are the highest among the three classifiers.

*5.5. Other Short-Circuit Fault Analysis.* The samples of PPF, DLGF, and TLGF are trained and classified by the same classification model. The training process of CNN is shown in Figure 8, and the classification results are shown in Table 3.

It can be seen that, when the voltage sag sources caused by different fault types are classified based on single transient characteristics, the classification accuracy of voltage sag sources caused by different fault types is seriously affected by different feature selections, different classifiers, and different fault types. The classification accuracy of KNN and EL is improved based on multidimensional features, but it cannot meet the requirements. Regardless of the voltage sag events

caused by any type of fault, CNN can be used to accurately classify based on the multidimensional feature matrix, and the classification accuracy is up to 100%.

*5.6. Comparison with Existing Methods.* The proposed method is based on multipoint monitoring data and can be used to identify voltage sag sources in loop and radiation networks. The existing voltage sag source positioning methods based on multimonitoring points include the branch current deviation method and node voltage deviation method. Among them, the node voltage deviation method does not consider SLGF, DLGF, and TLGF faults. Therefore, the proposed method is compared with the branch current deviation method.

The branch current deviation method defines the current outflow node as the positive direction. The deviation current

TABLE 4: Branch current deviation table.

Line	Node	Branch current deviation	Node	Branch current deviation
$L1_2$	1	A: 26.4884	2	A: 20.5416
		B: -0.0009		B: 0.0013
		C: 0.0033		C: -0.0028
$L1_5$	1	A: 0.0007	5	A: 0.0007
		B: 0.0004		B: 0.0004
		C: 0.0000		C: 0.0000
...	...	...	...	...
$L14_{13}$	13	A: 0.0000	14	A: 0.0000
		B: 0.0002		B: 0.0002
		C: 0.0003		C: 0.0003

TABLE 5: Location result.

Fault type	SLGF	PPF	DLGF	TLGF
Accuracy	48.24%	60.59%	48.24%	48.24%

of the branch is calculated by Equation (3), and the priority is determined according to the magnitude of the deviation current, thus determining the direction of the voltage sag source. The primary SLGF on line  $L1_2$  is positioned, and the branch current deviation method is briefly described as an example.

The deviation coefficients of all branch currents are calculated, as shown in Table 4. It can be seen that, according to the priority, fault current flows through line  $L1_2$  from node 1 and through line  $L1_2$  from node 2. Therefore, it can be judged that SLGF occurs on line  $L1_2$ .

Based on the branch current deviation method, the samples used for the CNN test are located for the voltage sag source, and the positioning results are shown in Table 5.

It can be seen from Table 5 that the branch current deviation method is used to locate the voltage sag caused by SLGF, DLGF, and TLGF, and the positioning accuracy is 48.24%. The accuracy of voltage sag positioning caused by PPF is 60.29%, so there is a certain gap with the method proposed in this paper. Based on the analysis of the positioning results, this method is greatly affected by the position of the power source, and the fault near the power source generates a larger fault current, so it can be positioned more accurately. Besides, the fault current in the neutral arc suppression coil system is composed of the whole system capacitive current, so this method is not applicable to the neutral arc suppression coil grounded system.

## 6. Conclusion

In summary, the voltage sag source location method based on the multidimensional characteristic matrix proposed in this paper can accurately locate the voltage sag caused by different fault types. Compared with the traditional voltage sag source location method based on multipoint monitoring data, it has

higher accuracy and is more suitable for the voltage sag source location in the loop network.

In this paper, a locating method of the voltage sag source based on the multidimensional feature matrix was proposed; then, a multidimensional feature matrix was constructed by extracting the disturbance power and disturbance energy, the real part of equivalent impedance, the slope of system trajectory, and the real part of the current at both ends of the whole network line, and then, CNN was used to realize the voltage sag source. The IEEE14 node model was built in PSCAD\EMTDC, and a large number of waveform data obtained by simulation under different fault types was adopted to verify the proposed method in this paper, and then, it was compared with the branch current deviation method. The simulation results show that the proposed method could accurately identify the location of voltage sag sources under different fault types, and the accuracy was higher than that of traditional methods.

## Conflicts of Interest

The authors declare no conflict of interest.

## Acknowledgments

This paper is sponsored by State Grid Science and Technology Project (Grant No. 52153218000B).

## References

- [1] IEEE Std, *Recommended practice for monitoring electric power quality*, Technical report, Draft 5, 1994.
- [2] Z. Yang, Z. Pingping, Y. Haojun, and L. Ganyun, "Review on cause analysis and source location for voltage sag," *Proceedings of the CSU-EPSA*, vol. 26, 2014no. 12, pp. 15–20, 2014.
- [3] Z. Yan, Y. Lin, and S. Zhenguo, "Multiobjective optimal allocation of monitors for voltage sag location under observability constraint," *Transactions of China Electrotechnical Society*, vol. 34, no. 11, pp. 2375–2383, 2019.
- [4] L. Xialin, L. Yajuan, and W. Zhu, "A new method to classify and identify composite voltage sag sources in distribution network," *Power System Protection & Control*, vol. 45, no. 2, pp. 131–139, 2017.
- [5] L. Ganyun, J. Xiaowei, H. Sipeng et al., "Location of voltage sag source based on semi-supervised SVM," *Power System Protection and Control*, vol. 47, no. 18, pp. 76–81, 2019.
- [6] A. C. Parsons, W. M. Grady, E. J. Powers, and J. C. Soward, "A direction finder for power quality disturbances based upon disturbance power and energy," *IEEE Transactions on Power Delivery*, vol. 15, no. 3, pp. 1081–1086, 2000.
- [7] T. Tayjasantant, C. Li, and W. Xu, "A resistance sign-based method for voltage sag source detection," *IEEE Transactions on Power Delivery*, vol. 20, no. 4, pp. 2544–2551, 2005.
- [8] C. Li, T. Tayjasantant, W. Xu, and X. Liu, "Method for voltage-sag-source detection by investigating slope of the system trajectory," *IEE Proceedings - Generation, Transmission and Distribution*, vol. 150, no. 3, pp. 367–372, 2003.
- [9] J. Gao, Q. Z. Li, and J. Wang, "Method for voltage sag disturbance source location by the real current component," in *Asia-Pacific Power and Energy Engineering Conference*, IEEE Computer Society Washington, DC, USA, 2011.

- [10] G. W. Chang, J.-P. Chao, H. M. Huang et al. et al., "On tracking the source location of voltage sags and utility shunt capacitor switching transients," *IEEE Transactions on Power Delivery*, vol. 23, no. 4, pp. 2124–2131, 2008.
- [11] B. Wang, W. Xu, and Z. Pan, "Voltage sag state estimation for power distribution systems," *IEEE Transactions on Power Systems*, vol. 20, no. 2, pp. 806–812, 2005.
- [12] H. WG, D. RA, X. Y. Zhang, R. Lu, G. QQ, and X. YH, "Comparison of methods for voltage sag source detection in distribution system," *Electrical Measurement & Instrumentation*, vol. 48, no. 8, pp. 53–58, 2011.
- [13] L. Haichao, O. Sen, and Z. Huaying, "A novel location method of power grid voltage sag source with BP neural network," *Journal of Electric Power Science and Technology*, vol. 2, 2017.
- [14] L. Ganyun and Y. Wu, "Optimization comprehensive criterion for voltage sag source location," *Power System Protection and Control*, vol. 5, pp. 66–71, 2013.
- [15] S. Weimeng, *Influence Assessment and Location Research of Voltage Sags Based Intelligent Algorithm*, Zhejiang Normal University, 2011.
- [16] K. Wang, *Research on Voltage Sag Source Location Method Based on Comprehensive Criterion*, China University of Mining and Technology, 2018.
- [17] L. Yingying, T. Wang, and F. Dandan, "Fault source localization method based on multiple criteria for voltage dips," *Proceedings of the CSEE*, vol. 35, no. 1, pp. 103–111, 2015.
- [18] Z. Chenyu, S. Mingming, F. Zhong, J. Zheng, and Y. Xiaodong, "Research on voltage sag event classification and short circuit type identification," *Electric Power Engineering Technology*, vol. 37, no. 2, pp. 102–107, 2018.
- [19] Y. Baocai, W. Wang, and L. Wang, "Review of deep learning," *Journal of Beijing University of Technology*, vol. 1, pp. 48–59, 2015.
- [20] C. Zheng, X. Kong, G. Jiahua et al., "Traveling wave head detection algorithm based on wavelet multiscale information fusion," *Smart Power*, vol. 47, no. 5, pp. 97–102, 2019.
- [21] C. Wei, J. He, and P. Xiping, "Classification for power quality disturbance based on phase-space reconstruction and convolution neural network," *Power System Protection and Control*, vol. 46, no. 14, pp. 87–93, 2018.
- [22] Y. Zhao, H. Li, S. Wan et al., "Knowledge-aided convolutional neural network for small organ segmentation," *IEEE Journal of Biomedical and Health Informatics*, vol. 23, no. 4, pp. 1363–1373, 2019.
- [23] L. Li, T. T. Goh, and D. Jin, "How textual quality of online reviews affect classification performance: a case of deep learning sentiment analysis," *Neural Computing and Applications*, vol. 32, no. 9, pp. 4387–4415, 2020.
- [24] Y. Zhang, J. Pan, L. Qi, and Q. He, "Privacy-preserving quality prediction for edge-based IoT services," *Future Generation Computer Systems*, vol. 114, pp. 336–348, 2021.
- [25] T.-T. Goh, Z. Xin, and D. Jin, "Habit formation in social media consumption: a case of political engagement," *Behaviour & Information Technology*, vol. 38, no. 3, pp. 273–288, 2019.
- [26] D. Jin, S. Shi, Y. Zhang, H. Abbas, and T.-T. Goh, "A complex event processing framework for an adaptive language learning system," *Future Generation Computer Systems*, vol. 92, pp. 857–867, 2019.
- [27] G. Jike, C. Dong, W. Wang, C. Zuqin, C. Guorong, and L. Can, "On the fire-disaster classification based on the improved naive Bayesian algorithm," *Journal of Safety and Environment*, vol. 19, no. 4, pp. 1122–1127, 2019.
- [28] S. K. Goudos, P. D. Diamantoulakis, and G. K. Karagiannidis, "Multi-objective optimization in 5G wireless networks with massive MIMO," *IEEE Communications Letters*, vol. 22, no. 11, pp. 2346–2349, 2018.
- [29] S. K. Goudos, Z. D. Zaharis, and K. B. Baltzis, "Particle swarm optimization as applied to electromagnetic design problems," *International Journal of Swarm Intelligence Research*, vol. 9, no. 2, pp. 47–82, 2018.
- [30] Z. Gao, H. Xue, and S. Wan, "Multiple discrimination and pairwise CNN for view-based 3D object retrieval," *Neural Networks*, vol. 125, pp. 290–302, 2020.
- [31] Y. Zhang, Y. Jin, J. Chen, S. Kan, Y. Cen, and Q. Cao, "PGAN: part-based nondirect coupling embedded GAN for person re-identification," *IEEE Multimedia*, vol. 27, no. 3, pp. 23–33, 2020.
- [32] S. Wan, X. Xu, T. Wang, and Z. Gu, "An intelligent video analysis method for abnormal event detection in intelligent transportation systems," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–9, 2020.
- [33] S. Ding, S. Qu, Y. Xi, and S. Wan, "Stimulus-driven and concept-driven analysis for image caption generation," *Neurocomputing*, vol. 398, pp. 520–530, 2020.
- [34] J. Jia, Q. Ruan, Y. Jin, G. An, and S. Ge, "View-specific subspace learning and re-ranking for semi-supervised person re-identification," *Pattern Recognition*, vol. 108, article 107568, 2020.
- [35] Z. Gao, Y. Li, and S. Wan, "Exploring deep learning for view-based 3D model retrieval," *ACM Transactions on Multimedia Computing, Communications, and Applications*, vol. 16, no. 1, pp. 1–21, 2020.
- [36] Y. Xi, Y. Zhang, S. Ding, and S. Wan, "Visual question answering model based on visual relationship detection," *Signal Processing: Image Communication*, vol. 80, p. 115648, 2020.