

Intermittent Arc Fault Detection Based on Machine Learning in Resonant Grounding Distribution Systems

Ye Tian¹, Mou-Fa Guo², *Member, IEEE, Member, CSEE*, and Duan-Yu Chen, *Member, IEEE*

Abstract—In resonant grounding systems, most single-phase-to-ground faults evolve from IAFs (Intermittent Arc Faults). Earlier detection of IAFs can facilitate fault avoidance. This work proposes a novel method based on machine learning for detecting IAFs in three steps. First, the feature of zero-sequence current is automatically extracted and selected by a newly-designed FINET (“For IAFs, Neuron Elaboration Net”), instead of traditional feature selection based on time-frequency decomposition. Moreover, data of the zero-sequence current divided by different time windows are successively input into the trained FINET. A proposed PSF (principal-subordinate factor) analyses the results obtained from FINET to improve anti-interference in the mentioned IAF detection algorithm. Experiments using PSCAD/EMTDC software simulation data show the proposed method is feasible and highly adaptable. In addition, the detection result of on-site recorded data demonstrates the effectiveness of the proposed method in practical resonant grounding systems.

Index Terms—Resonant grounding distribution systems, intermittent arc faults (IAFs) detection, “For IAFs, Neuron Elaboration Net” (FINET), principal-subordinate factor (PSF).

I. INTRODUCTION

AGING insulation or contact between a power transition line and a high impedance surface (e.g., trees, gravel, concrete, turf or asphalt pavement) will cause the impedance of transition to change based on a periodic arc, and thus potentially result in IAFs [1]. Intermittent arc grounding exceeding relative voltage will cause a rapid change in the grid operating state, further causing high-frequency oscillation to exceed relative voltage [2]. IAFs can cause significant damage to electric power systems. However, the amplitude of the fault current is insufficiently high to trigger normal power system

supervision. Therefore, the detection of IAFs is still a topic worth studying.

An exhaustive study of general power system fault detection has been well documented in [3]. Unlike general power system faults, IAFs are difficult to maintain in a continuous state. Observable sinusoidal fault signals indicate that IAFs don't follow a clear pattern, which makes them more difficult to detect. Prevalent methods for IAF detection are a combination strategy of feature extraction and fault decision-making [4]. Several representative methods have been developed for hand-crafted feature extraction including multiresolution morphological gradient (MMG) [5], wavelet transform (WT) [6], S transform (ST) [7], mathematical morphology [8], adaptive Kalman filter (AKF) [9], and advanced distortion detection techniques (ADDT) [10]. In terms of fault decision-making methods, artificial neural networks (ANN) [11], fuzzy inference systems (FIS) [12], and support vector machines (SVM) [13] or multiple SVMs [14] are all widely used.

In [5], MMG is used to extract current waveform characteristics of 3.5 cycles post-fault, and these features are then processed using a multi-layer perceptron neural network. In [15], WT is used to decompose the fault waveforms and extract features, which are then input to ANN for identification. The above strategies require accurate fault transient waveform data. On the one hand, this makes manual interception of fault waveforms a lot of work; on the other hand, uncertain noise in the actual environment will affect the effectiveness of different strategies. In [16], empirical mode decomposition is used to extract the frequency features of long-cycle fault signals, which could then be further recognized by ANN. However, this method requires longer waveform data, making it difficult to detect IAFs in real-time. In addition, the intermittent IAF process may cause the algorithm not to work.

The complexity of IAFs makes it difficult for existing strategies to classify short-cycle or long-cycle waveforms. Development of machine learning technology has allowed manual extraction of mathematical features to be replaced by neural networks. Current machine learning methods have more computing nodes and branches. Each node can use more efficient calculation methods such as convolution operation to extract features, and these features are transferred to the end node in the form of jumping nodes, greatly enhancing the learning ability of the machine and realizing effective recognition of the trained samples. In addition, it allows for

Manuscript received December 25, 2019; revised August 15, 2020; accepted September 21, 2020. Date of online publication November 20, 2020; date of current version November 25, 2021. This work was sponsored by the National Natural Science Foundation of China (No.51677030).

Y. Tian (ORCID: <https://orcid.org/0000-0002-8924-6220>) is with the College of Electrical Engineering and Automation, Fuzhou University, Fuzhou 350108, China; the Department of Electrical Engineering, Yuan Ze University, Taiwan 32003, China; and also with the Fuzhou Power Supply Company of State Grid Fujian Electric Power Co., Ltd., Fuzhou 350009, China.

M. F. Guo (corresponding author, email: gmf@fzu.edu.cn; ORCID: <http://orcid.org/0000-0002-2358-0509>) is with the College of Electrical Engineering and Automation, Fuzhou University, Fuzhou 350108, China.

D. Y. Chen is with the Department of Electrical Engineering, Yuan Ze University, Taiwan 32003, China.

DOI: 10.17775/CSEEJPES.2019.03410

reliable judgements for unknown samples. The optimal feature quantity of short-cycle fault waveform can be automatically self-summarized and extracted by the machine, and compared with conventional classifiers, the density of neuron nodes in neural networks allows for more comprehensive extraction of features. Furthermore, neural network parameters can be trained to obtain an optimal fit.

Machine learning uses multiple neural network architectures, including perceptron, deep confidence networks, convolutional neural networks (CNNs), and recurrent neural networks (RNNs). Among these, RNNs are suitable for waveform recognition with causality while CNNs are ideal for two-dimensional matrix recognition with multiple local features. In [17], CNNs are used to identify the spectrum map of the distribution network fault. Since CNNs provide strong image processing power, the recognition rate of image content improved significantly. Although machine learning has performed well in areas such as image processing [18], the detection of faults in distribution networks is still in its infancy. Earlier research asked human subjects to actively describe certain features to train the network model, leaving neural networks unable to self-search or conclude rules from known empirical data. At the same time, since the actual grid cannot perform fault experiments, the number of actual fault samples recorded is limited. Hence, most methods use software simulation data as training samples and actual data as testing samples. The use of transient process characteristics of one-time simulation fault data as training samples raises two problems. First of all, the one-time simulation waveform is not fully utilized; secondly, if the fault transient process is disturbed, the classification capability of conventional classifiers will be limited. In other words, simple classifiers cannot avoid over-fitting problems due to insufficient data or diversity [19]. Deep neural network architectures provide some solutions to avoid over-fitting, but these often result in gradient disappearance or gradient explosion [20]. Nowadays, due to rapid development of artificial intelligence (AI) technology, multi-branch, refined neural networks with inception structures hold out some prospect for a solution to this problem [21]. To a certain degree, this kind of network architecture improves the limitations of traditional neural networks [22].

Traditional methods suffer from two main problems. One is that previous strategies used only one segmentation in the fault signal as the detection object, ignoring the intermittent process of IAFs; the other is the training of traditional neural networks is limited by the number of samples, which will produce a problem of gradient disappearance or over-fitting. In order to achieve effective detection of short-cycle waveform signals, this study proposes FINET, a neural network architecture optimization designed to identify IAFs in resonant grounding distribution networks. The zero-sequence current waveform at the outgoing point of the main substation is used to identify IAFs on the feeder. The waveform is divided into several detection windows at different times (including fault transient and steady state processes), and then become the FINET input. These continuous windows are identified by FINET as IAFs and non-IAFs (normal, load switching, etc.) and further input to PSF. The PSF method is composed of principal factor

(PF) counter and subordinate factor (SF) counter. FINET outputs are jointly counted by the two counters in PSF. When required conditions of the PSF method are satisfied, the occurrence of IAFs is confirmed. Since the final decision is mainly based on computation results, the possibility of fault misjudgment is significantly reduced. To study performance of the proposed scheme, different distribution network architectures and various events are simulated in power systems computer aided design/electromagnetic transients including DC (PSCAD/EMTDC). In addition to simulation data, the measured data (hardware experimental data and field data) are also adopted for analysis.

The main contributions of this paper are summarized as follows:

- By dividing the fault waveform, problems of small number of training samples and non-representative sample characteristics are solved.
- A novel neural network architecture (FINET) is proposed to solve the problems of neural network over-fitting and gradient disappearance during training.
- Since the traditional method has few criteria, the problem of missed detection or misjudgment of the measured waveform is solved by inputting the segmented zero-sequence current waveform to the proposed detection method (FINET+PSF).

The rest of this paper is organized as follows: Section II provides detailed information on the resonant grounding distribution system model, events model, simulation and measured data. Section III introduces the optimal neural network architecture for IAFs, and provides a scheme using FINET+PSF to implement IAFs detection. Section IV compares the proposed detection methods with other established approaches in simulation and engineering practice; Section V draws some concluding remarks.

II. MODELING AND SIMULATION

Different kinds of events (IAFs, single-phase-to-ground fault, load switching and capacitor bank switching) in resonant grounding distribution systems are simulated using PSCAD/EMTDC software. Typical radial and branch networks are respectively depicted in Fig. 1(a) and Fig. 1(b). Figure 1(c) simulates an actual 10 kV resonant grounding distribution network of China Fujian Power Co., Ltd. OL and CL denote the length of the overhead and cable line, respectively. The above systems are connected to an ideal power supply, the main transformer has a capacity of 20 MVA, and the main line parameters of the simulation model are listed in Table I.

TABLE I
LINE PARAMETERS

Line types	Phase-sequence	R (Ω/km)	L (mH/km)	C ($\mu\text{F}/\text{km}$)
Overhead line	Positive-sequence	0.17	1.2	0.0096
	Zero-sequence	0.23	5.5	0.0080
Cable line	Positive-sequence	0.27	0.255	0.339
	Zero-sequence	2.7	1.102	0.28

IAFs are difficult to simulate and identify because of the following characteristics [4]:

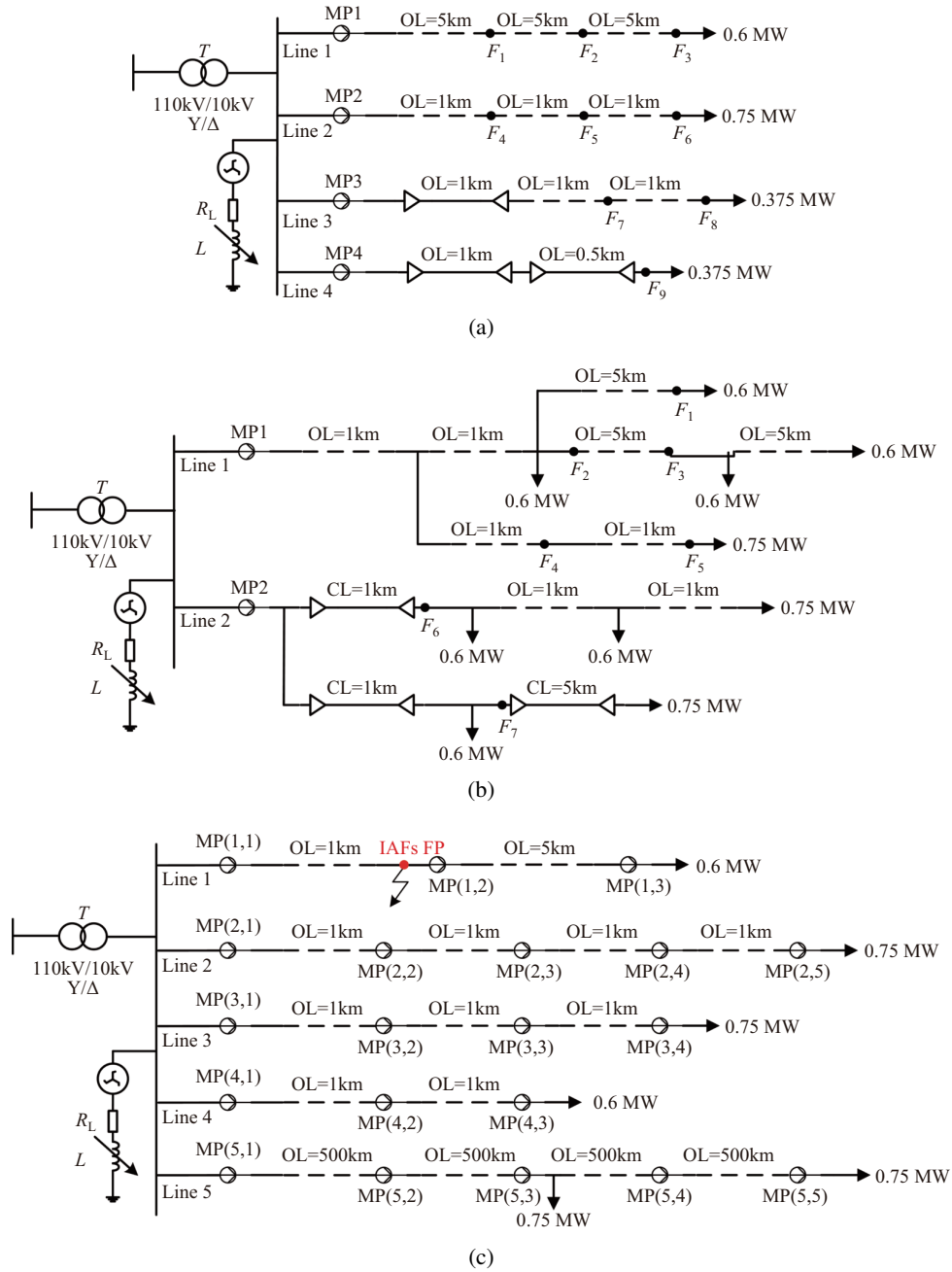


Fig. 1. Software simulation models. (a) Model 1: Radial network. (b) Model 2: Branch network. (c) Model 3: Field network.

- 1) Weak fault current: The value of current at the fault location is less than 10% of the total system load.
- 2) Nonlinearity: Because of the existence of fault arc and grounding medium, the fault current presents non-linearity.
- 3) Randomness: Arc instability creates randomness in the current.
- 4) Intermittence: Arc burning and extinguishing result in current intermittency.
- 5) Asymmetry of current waveform: The positive and negative half-cycle waves of the fault current are not completely symmetrical.

Most previous arc models simulate arc characteristics based on high-frequency arc extinguishing theory or power frequency

arc extinguishing theory. Common arc models include those proposed by Cassie, Mayr, Schwarz, cybernetics, etc., where the cybernetic model is suitable for simulating small current arc faults [23]. Since the distribution network mostly uses low current grounding, the IAF simulation in this study uses a cybernetic model.

The mathematical expression of cybernetics is as follows:

$$\begin{cases} \frac{dg_{\text{arc}}}{dt} = \frac{1}{\tau_c}(G_c - g_{\text{arc}}) \\ G_c = \frac{|i|}{V_c \times L_c}, \tau_c = \beta \frac{I_c}{L_c} \end{cases} \quad (1)$$

where G_c is steady-state conductance, and τ_c is arc time constant, V_c is the steady-state field strength in the arc column,

which is approximately constant. The empirical value $V_c = 15$ V/cm, L_c is the arc length in cm. The coefficient β takes the empirical value 2.85×10^{-5} . I_c is the peak value of arc current, which is approximate to the short-circuit current value of direct grounding in kA.

As shown in Fig. 2(a), the fault model is connected in series by a cybernetic resistance (R_{arc}) and a random resistance (R_r). The model is a single-phase-to-ground fault model when $R_{arc} = 0$ and R_r is constant, and a IAFs model when R_{arc} is controlled by cybernetics and R_r varies with transient moments. As shown in Fig. 2(b) and Fig. 2(c), in the load switching model, the resistive three-phase unbalanced load accounts for 10% ~ 25% of the total system load. In the capacitor bank switching model, the reactive power of the shunt compensation capacitor bank is 3600 kvar. These four events are simulated at each fault location of model 1 (F1 ~ F9) and model 2 (F1 ~ F7). In addition to software simulation data, we use both the experimental data of the hardware system [24], [25] and field data provided by Fujian Electric Power Co., Ltd. Field data is randomly selected from different distribution network architectures and different lines. As shown in Fig. 3(a), an IAF experiment was carried out on the hardware experimental system, with a recording device [26] recording the experimental waveform. All the above mentioned data are summarized in Table II.

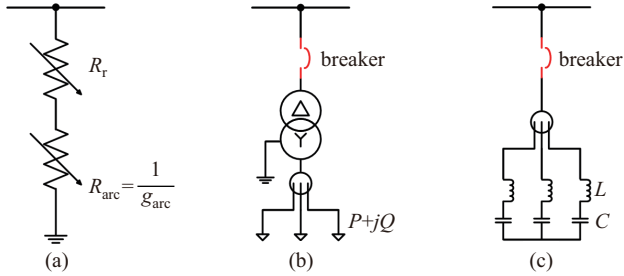


Fig. 2. Events model. (a) IAFs/single-phase-to-ground fault. (b) PQ-load switching. (c) Capacitor bank switching.

TABLE II
COMPLETE WAVEFORMS INFORMATION

Waveform type	Number of complete zero-sequence current waveforms			
	Software simulation		Measured	
	Model 1	Model 2	Hardware experiment	Field data
Intermittent arc faults	162	108	12	10
Single-phase-to-ground fault	270	180	54	11
Load switching	54	36	NA	NA
Capacitor bank switching	32	23	NA	NA
Total(IAFs-non-IAFs)	162–356	108–239	12–54	10–11

As shown in Fig. 3(b), FINET inference on the edge side uses a Jetson TX2, which integrates the GPU and CPU. The zero-sequence current waveform data obtained by the recording devices is input to TX2 for recognition through RS 485 communication. TX2 detects whether IAFs occur in real time and stores detection results in memory. The detection results are sent to a portable device such as a smart-phone or a notebook computer via HTTP.

III. INTERMITTENT ARC FAULTS DETECTION METHOD

The waveform of IAFs changes continuously with severity of the fault. Hence, it is necessary to pay attention to its intermittent and imprecise changes during the detection process. In addition, a faster recognition speed needs to be considered during the detection process. Since single-phase-to-ground faults usually evolve from IAFs, real-time identification of IAFs can effectively prevent more serious faults from occurring.

Traditional machine learning methods include SVM, Adaboost, and so on. Early on, these methods provided good results for simple regression problems. However, with increasing difficulty of the classification problem and limitations of computing nodes, network structures with more layers such as RNN and CNN have emerged for classification. The fault characteristics of IAFs are not obvious. Although the

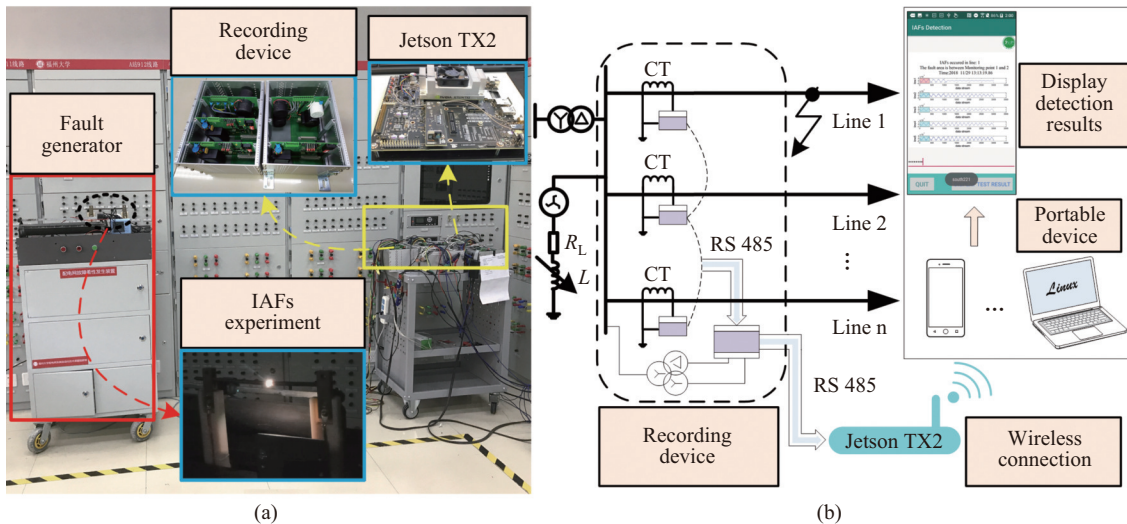


Fig. 3. Hardware condition. (a) Hardware experimental system. (b) Hardware detection system structure diagram.

fault waveform “zero rest” and non-linear characteristics are retained during the segmentation process, this brings higher requirements to the machine learning method. In addition to sequential neural networks such as RNN and chain-CNN, multi-branch and dense-connected neural networks can have stronger learning capabilities. Not only can they accurately identify trained data sets, but they can also effectively classify data samples with similar characteristics.

The traditional IAF detection methods only intercept some features of a single fault waveform to make a judgment and ignore the intermittent processing of IAFs. The mathematical feature extraction method of artificial designs usually involve too many parameters, making it difficult to optimize the method, and it is more difficult to set the fault threshold in practice. In fact, the richer the extracted feature information, the better the classification effect. The selection, combination, and optimization of waveform feature values require a lot of practical experience and research work, and machine learning is a continuous trial and error optimization system through the computer until the output result of neural networks is accurate. It is very difficult to achieve this effect with manual design methods to replace this process. Compared with common IAF detection methods, machine learning methods are more intelligent. Through automatic iterative processes, the network can be trained to obtain the best classification effect. Because the convolutional network has translation invariance and a spatial hierarchical structure of modes, the shape features of the IAF waveform can be effectively extracted, also avoiding the problem of difficult threshold settings.

The overall experiment is shown in Fig. 4 with three steps. Step 1: the zero-sequence current waveform of software simulation model 1 is processed into training samples and input into FINET for training. Finally, the trained FINET is obtained. Step 2: the waveform of software simulation model 2, the experimental waveform of hardware system and field waveform are processed into testing samples and the performance of FINET and other neural networks were compared. Step 3: trained FINET and PSF methods are combined to implement real-time IAFs detection and final detection accuracy is compared with other state-of-the-art methods.

A. Acquisition of Training and Testing Samples

As shown in Fig. 5(b) and Fig. 5(c), detection windows

(Win(k)) are manually intercepted and the length of each window is 2 power frequency periods. Due to the periodicity of the fault process, each window basically retains the IAFs waveform non-linearity, non-stationary and other characteristics, which includes two similar peaks and troughs. The number of detection windows is shown in Table III. The sampling frequency (f_{Recorder}) of the software simulation and hardware experiment is 10 kHz, so the first-order tensor length of each detection window is 400 ($2 \times 10 \text{ k} \times \frac{1}{50}$). Similarly, the field data sampling frequency is 4 kHz and the tensor length is 160.

TABLE III
THE NUMBER OF DETECTION WINDOWS

Data source		Number of detection windows			Fault point	Monitoring point
		IAF	non-IAF	Total		
Software simulation	Model 1	1082	2497	3579	F1 ~ F9	MP1 ~ MP4
	Model 2	656	1597	2253	F1 ~ F7	MP1 ~ MP2
Measured	Hardware experiment	52	156	208	OMIT	OMIT
	Field data	41	66	107	OMIT	OMIT

As shown in Fig. 5(d), to standardize data at different sampling frequencies [$f_{\text{Recorder}} \in (0, 25.6) \text{ kHz}$], pretreatment process is as follows:

First, as shown in (2), we use the bicubic interpolation in open source computer vision library (OpenCV) to map the detection window's tensor length to 1024.

$$g(x) = \sum_{i=0}^3 a_i x^i \quad (2)$$

where $g(x)$ is the interpolation function, a_i is the interpolation coefficient, x is interpolation interval elements.

Second, as shown in (3), the min-max normalization method is used to normalize the value of each detection window's element to between 0 and 1,

$$x_i = \frac{x_i - \min(\text{Win}(k))}{\max(\text{Win}(k)) - \min(\text{Win}(k))} \quad (3)$$

where x_i is the window element, $\text{Win}(k)$ is the detection window, and the $\max(\cdot)$ and $\min(\cdot)$ are respectively the maximum and minimum operators.

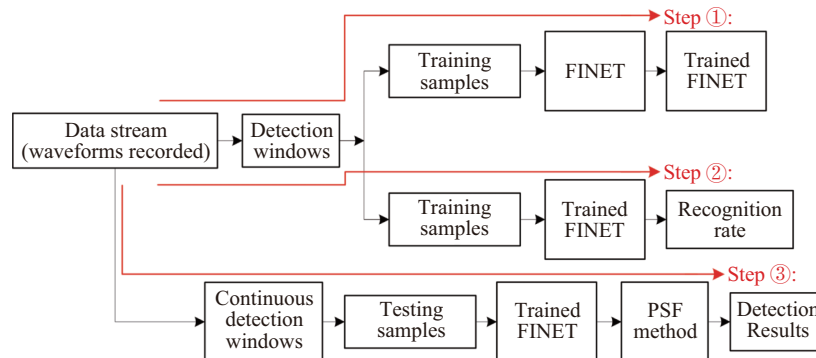


Fig. 4. Flow chart of the proposed IAFs detection in resonant grounding distribution systems.

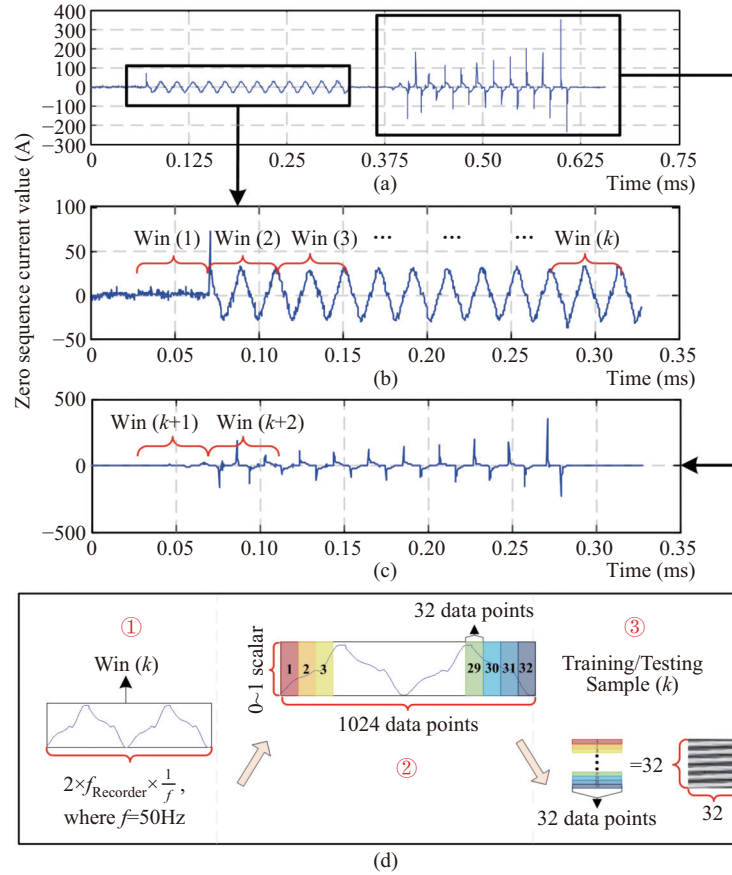


Fig. 5. Zero-sequence current waveforms. (a) IAFs wave. (b) IAFs wave (with high transition impedance). (c) IAFs wave (with low transition impedance). (d) Pretreatment process.

Finally, with the change of tensor of the window data, 1024 data points are divided into 32 segments, with 32 data points in each segment. Therefore, a sample for training or testing is formed in terms of a 32×32 square matrix. The method solves the sampling frequency mismatch problem, and calculation is efficient and robust.

B. Design and Training Process for FINET

Given Alexnet's successful application of CNN in image recognition in 2012 [27], mainstream work now focuses on deepening neural networks. Usually, the "depth" in a deep network refers to a series of continuous presentation layers; how many layers are contained in the data model, which is called the depth of the model. Other data representations that only learn one or two layers are called shallow learning. But an excessive number of network layers will cause two limitations: First, too many parameters would result in application difficulty and over-fitting due to insufficient training samples. Second, the gradient disappears. In practice, it is found that recognition accuracy cannot be increased using a simple sequential CNN structure, and calculation speed is slow. In [21], an inception structure is proposed to reduce the number of parameters by using sparse layer connections instead of sequential layer connections. This structure is more conducive to GPU calculation of dense matrices to speed up the calculation process. Although deepening the network

model will enhance its learning ability, an excessively deep network model will complicate parameter optimization and produce degradation through accuracy saturation. In [28], a deep network architecture called DenseNet is proposed, which creates a leaping connection method so input of the latter layer is the output of all the previous layers. The use of dense blocks associates layers with layers to avoid gradient disappearance. The fitting effect of this kind of network works well, but the structure of dense networks produces excessive numbers of the input and model parameters, reducing calculation speed. According to the above two methods, this paper innovatively proposes a structure in which the inception structure and the Dense block are merged to correlate the layers. This method not only improves calculation speed (the network parameters are less than 10% of these deep networks) but also has good network fitting effect, reducing the problem of gradient disappearance. Complete architecture of FINET is shown in Fig. 6. The network only uses part of the blocks of the above-mentioned deep network architecture, and the number of layers is much smaller. In identified results, "1" indicates that FINET recognizes the detection window as IAF, and "0" indicates that it is non-IAF. In the 16-layer FINET network, the structure between the 2nd ~ 11th layers is a fusion of inception and Dense blocks, which replaces the artificial design to extract features. The 12th ~ 15th are fully connected layers used as the classifier as follows:

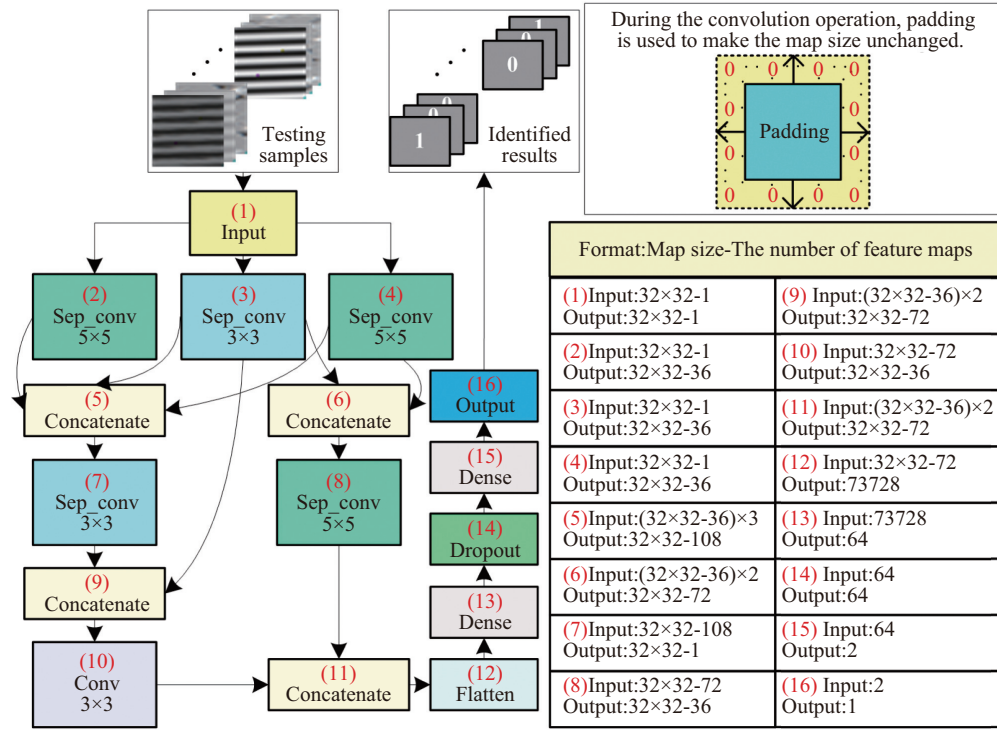


Fig. 6. The FINET architecture designed for IAF.

- 1) Taking the 3rd layer as an example, as shown in Fig. 7(a), the separable convolution consists of two stages. First, a single 3×3 convolution kernel is used for feature extraction while zero-padding is applied while maintaining the size of the output feature maps. In the second stage, 36 one-dimensional convolution kernels are used to generate 36 feature maps. Separable convolution reduces kernel parameters by more than 60% compared to traditional convolution.
- 2) Taking the 11th layer as an example, as shown in Fig. 7(b), the concatenate layer combines feature maps outputted by the 8th and 10th layers into 72 feature maps.

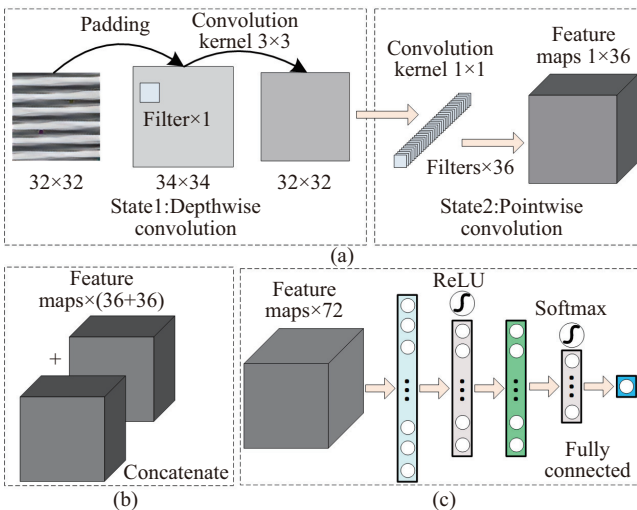


Fig. 7. Layer operation. (a) Separable convolution layer. (b) Concatenate layer. (c) Fully connected layer.

- 3) The 12th to 15th layers are a fully connected layer. As shown in Fig. 7(c), the 12th layer expands the feature map into a first-order tensor, the activation function of 13th layer is “ReLU” [29], and the 14th layer drop 25% of features to prevent over-fitting. In addition, the activation function of the 15th layer is “Softmax” [30] and the loss function is “categorical cross-entropy” [30].

There are 3579 detection windows in the simulated radial network, 90% are randomly selected as a training set and the rest are used as the validation set. The structure of the fully connected layer remains unchanged, and the connection mode of the 2nd ~ 11th layer is optimized until the FINET structure with the fastest convergence rate and the highest accuracy is obtained. Fig. 8 shows performance in terms of validation accuracy and loss value during the training process. Therefore, in the experiment, the model is selected after training in 23 epochs.

Evaluating the error of a machine learning model consists of 3 parts: model error = variance + bias + random error of noise.

Due to random errors from noise, there is no perfect predictive model. Hence, the smaller the variance and bias, the higher accuracy of the model.

The neural network model with few neurons will be under-fitting, both of variance and bias is relatively large; while the model with numerous neurons will be overfitting, the variance tends to be large, the bias is opposite. Face with samples that the model has never seen, the error ($|\text{predicted value} - \text{actual value}|$) is large. Therefore, reducing the variance of complicated neural networks as much as possible, that is, avoiding overfitting is the key to convergence of machine learning.

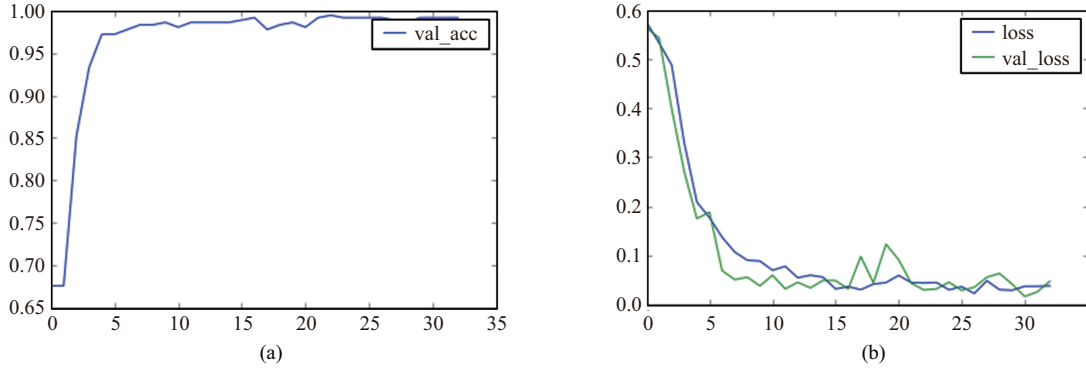


Fig. 8. Accuracy and loss of FINET. (a) Validation accuracy, (b) Training set and validation set loss.

If there is a large error in the prediction of the model, the model can be corrected through the following 5 points:

- 1) Add training data: Adding measured data to the training set can greatly reduce the variance.
- 2) Data augmentation: Artificially expand the amount of data and increase sample diversity to solve overfitting.
- 3) Adjust the dropout ratio: Make the weight of neuron nodes more uniform to avoid overfitting.
- 4) Batch normalization: Mix the simulated and measured data sets, divide them into several sets, normalize them separately, and train the processed data to avoid overfitting.
- 5) The model stops training as soon as it begins to degrade in performance on the validation set.

C. Principal-subordinate Factor Method

The generation of IAFs due to insulation equipment aging is simulated at the fault point (IAFs FP) of line 1 and the zero-sequence current waveform is recorded at the monitoring point [MP(1, 1)] in Fig. 1(c). As shown in Fig. 9(a), it is a curve of the above simulation waveform.

The random behavior of IAFs is set as follows:

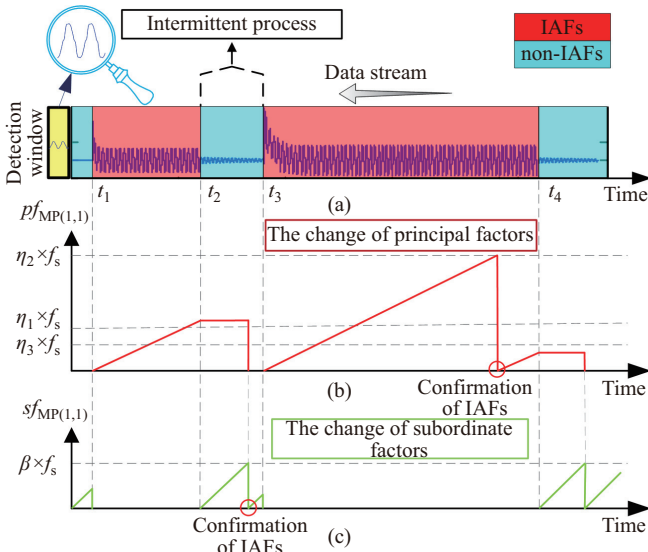


Fig. 9. Fault decision by PSF method. (a) Intermittent process of IAFs. (b) Ideal PF curve. (c) Ideal PF curve.

- 1) Simulation time:
 $|t2 - t1| \in [0.3, 0.8] \text{ s}$; $|t3 - t2| \in [0.1, 0.3] \text{ s}$;
 $|t4 - t3| \in [1.5, 2] \text{ s}$.
- 2) Simulation transition resistance:
 At $t1 \sim t2$, R_{arc} is controlled, $R_r \in [0.5, 2] \text{ k}\Omega$;
 At $t2 \sim t3$, $R_{arc} = 0 \Omega$, $R_r \in [10, 100] \text{ k}\Omega$;
 At $t3 \sim t4$, R_{arc} is controlled, $R_r \in [10, 300] \Omega$;
 After $t4$, $R_{arc} = 0\Omega$, $R_r > 100 \text{ k}\Omega$.

The random process can be described as follows:

- 1) At $t1$, the line changes from a healthy working state to a short-term IAF state.
- 2) At $t2$, the line changes to the intermittent state.
- 3) At $t3$, the line is re-ignited and converted to a long-term IAF state.
- 4) At $t4$, the IAFs are extinguished again, since the insulation is not completely broken down, the line temporarily returns to a healthy state.

As shown in Fig. 10, the complete zero-sequence current waveform is divided into sequential detection windows at different times according to the sampling step (f_s). FINET identifies them as IAF(1) or non-IAF(0), and then inputs the 0 or 1 sequence into the PSF method for statistical analysis. The decision made by the PSF method is the final detection

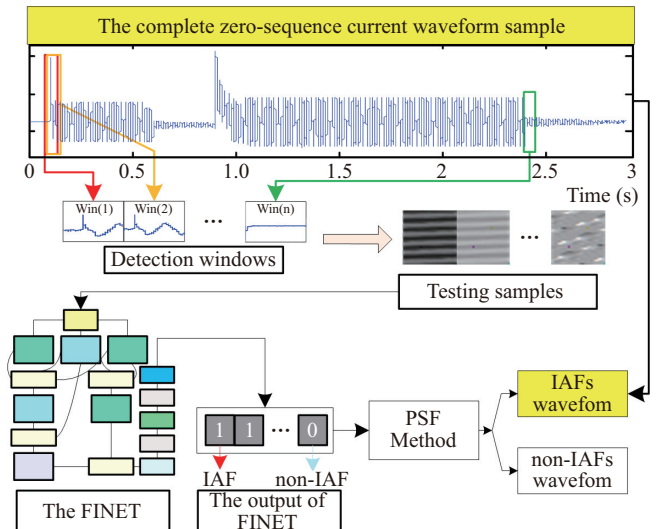


Fig. 10. The overall flow chart of FINET+PSF method.

result of complete zero-sequence current waveforms (Table II). According to the actual situation, the sampling step can be appropriately set ($f_s = 1000 \sim 5000$ windows·s⁻¹). The advantage is that doing so avoids redundant judgment by FINET and can implement real-time detection (the detection time of each window requires 216.695 μ s with a single Nvidia GTX 2080 GPU, and requires 1.278 ms on TX2).

Most scholars use mathematical extraction of fault features + a classifier to identify faults, but this method is limited by the ability of the classifier. The use of machine learning classifications and statistics to achieve fault waveform detection can greatly increase detection accuracy. The complete PSF method is shown in Algorithm 1, and the experimental parameters in PSF method are shown in Table IV. The ideal curves of PF and SF are respectively shown in Fig. 9(b) and (c), and the working principle of the PSF method is as follows:

- 1) When IAF does not occur, the two-cycle waveform in the detection window is identified as “0” by FINET, and SF = SF + 1.
- 2) When IAF occurs, the output of FINET is “1”, and PF = PF + 1.
- 3) When PF = PF + 1, SF is set as 0.
- 4) Short-term IAFs were confirmed as occurring when $PF > \eta_1 \times f_s$ and $SF > \beta \times f_s$ are satisfied.
- 5) Long-term IAFs were confirmed as occurring when $PF > \eta_2 \times f_s$ is satisfied.
- 6) When $PF < \eta_3 \times f_s$ and $SF > \beta \times f_s$ are satisfied, both PF and SF are set as 0.

TABLE IV
PARAMETER VALUE OR RANGE IN THE PSF METHOD

Parameter	η_1	β	η_2	η_3
Value/range	0.2 ~ 0.3	0.1	1 ~ 2	0.05

Traditional methods have difficulty overcoming interference caused by the IAF intermittent process. Through the dividing waveform, the complete fault process of IAFs is preserved. FINET continuously judges the process and outputs whether each divided waveform is a fault waveform. This transient, steady and intermittent waveform information is fully counted by the PSF method, thus improving both recognition accuracy and resistance to interference.

IV. RESULTS AND PRACTICE

A. Comparison with Other Methods

In Fig. 11, the structure of the chain-CNN is in the form of a convolution layers-fully connected layer. Similar to the chain-CNN, the structure of the RNN is in the form of an LSTM/GRU units-fully connected layer. Different from the traditional structure, FINET has more branches, and layers are not connected one by one. Besides, in addition to subsequent connections, FINET's neurons also have jump connections (as shown by the dotted line), so the proposed structure is smarter than a traditional structure. These four kinds of neural networks (chain-CNN, LSTM-RNN, GRU-RNN, and the proposed FINET) were trained using the same training samples and were tested using the same testing samples.

Algorithm 1: The algorithm of PSF method

Input: The real-time recognition results of continuous detection windows from FINET.

Output: The final detection results of IAFs.

- 1 Begin: Take model 3 in Fig. 1 as an example, define
- 2 PF =
$$\begin{pmatrix} pf_{MP(1,1)}pf_{MP(1,2)}pf_{MP(1,3)} \\ pf_{MP(2,1)}pf_{MP(2,2)}pf_{MP(2,3)}pf_{MP(2,4)}pf_{MP(2,5)} \\ pf_{MP(3,1)}pf_{MP(3,2)}pf_{MP(3,3)}pf_{MP(3,4)} \\ pf_{MP(4,1)}pf_{MP(4,2)}pf_{MP(4,3)} \\ pf_{MP(5,1)}pf_{MP(5,2)}pf_{MP(5,3)}pf_{MP(5,4)}pf_{MP(5,5)} \end{pmatrix},$$
- 3 SF =
$$\begin{pmatrix} sf_{MP(1,1)} \\ sf_{MP(2,1)} \\ sf_{MP(3,1)} \\ sf_{MP(4,1)} \\ sf_{MP(5,1)} \end{pmatrix}$$
- 4 Init: PF = 0, SF = 0
- 5 Step 1:
- 6 **if** the window (MP_(i,j), $j = 1, i = 1 \sim 5$) is IAF **then**
- 7 $pf_{MP(i,1)} \leftarrow pf_{MP(i,1)} + 1, sf_{MP(i,1)} \leftarrow 0$
- 8 **goto** Step 2
- 9 **else**
- 10 $sf_{MP(i,1)} \leftarrow sf_{MP(i,1)} + 1$
- 11 **goto** Step 3
- 12 **end**
- 13 Step 2:
- 14 **if** the window (MP_(i,j), $j > 1, i = 1 \sim 5$) is IAF **then**
- 15 $pf_{MP(i,j)} \leftarrow pf_{MP(i,j)} + 1$
- 16 **goto** Step 3
- 17 **else**
- 18 **goto** Step 3
- 19 **end**
- 20 Step 3:
- 21 $l \leftarrow \text{where}(pf_{MP(i,1)} = \max[pf_{MP(i,1)}])$
- 22 $t \leftarrow \text{where}(sf_{MP(i,1)} = \max[sf_{MP(i,1)}])$ {where(.) and max(.) are from the numpy function library}
- 23 **if** ($pf_{MP(l,1)} > \eta_1 \times f_s$ and $sf_{MP(l,1)} > \beta \times f_s$) or ($pf_{MP(l,1)} > \eta_2 \times f_s$)? **then**
- 24 IAFs may occur on line l
- 25 $p \leftarrow \text{where}(pf_{MP(l,j)} = \max[pf_{MP(l,j)}])$
- 26 IAFs may occur between MP(l, p) and MP($l, p + 1$)
- 27 **goto** Init
- 28 **else if** ($sf_{MP(t,1)} > \beta \times f_s$) and ($pf_{MP(t,1)} < \eta_3 \times f_s$) **then**
- 29 **goto** Init
- 30 **else**
- 31 **goto** Step 1
- 32 **end**

Software simulation data (model 1) are used as the training samples, while software simulation data (model 2), hardware experimental data and field data are used as testing samples (the number of samples is consistent with the number of detection windows in Table III). Test results are shown in Table V, and the test accuracy (φ) of the calculation formula is as shown in (4).

$$\varphi = \frac{\text{Correct numbers of detection windows}}{\text{Total numbers of detection windows}} \times 100\% \quad (4)$$

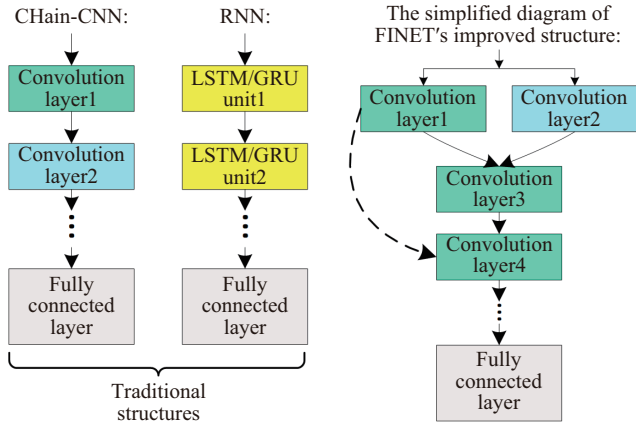


Fig. 11. Comparison between proposed structure and traditional structures.

TABLE V
RECOGNITION RATE OF FINET IN COMPARISON WITH OTHER NEURAL NETWORK ARCHITECTURES

Network architecture type	Testing accuracy (ϕ)		
	Software simulation (branch network)	Hardware experiment	Field
LSTM-RNN [31]	88.8%	81.7%	80.4%
GRU-RNN [32]	88.7%	82.2%	86.9%
Chain-CNN [17]	70.9%	75.0%	61.7%
FINET	94.9%	90.4%	94.4%

Table V shows that FINET designed for IAFs has higher accuracy for different distribution network architectures in terms of both simulation and measured data. Waveforms of IAFs have intermittent and distortion characteristics and change uncertainly with time, making them difficult to effectively identify. By dividing the IAF waveform, every two periodic waveforms are processed into a 32×32 matrix. Short time-series samples are not conducive to identification of LSTM or GRU type RNN, so recognition accuracy is not high. Although CNN is suitable for matrix classification, segmented samples also bring learning challenges. The learning ability and generalization ability of chain-CNN are not sufficient for effective identification. Compared with traditional RNN and CNN, FINET has more branches and nodes, and learning ability is enhanced while avoiding the problems of overfitting and gradient disappearance. By dividing the waveform, a small amount of measured data can also obtain a larger number of detection windows, proving the effectiveness of the method in terms of recognition rate.

As shown in Fig. 12, the information of feature maps extracted by FINET is significantly richer than those extracted through chain-CNN. As for LSTM/GRU-RNN, because there is less short-cycle waveform information after division, the traditional RNN structure cannot obtain enough fault features, so it cannot obtain higher recognition accuracy.

The training sets used in most other papers is composed of a combination of simulation and measured samples, while the remainder are tested to obtain high accuracy. However, the testing set samples in this paper have never been trained, hence, this is a more difficult task. Existing machine learning methods have not been completely correct in identifying

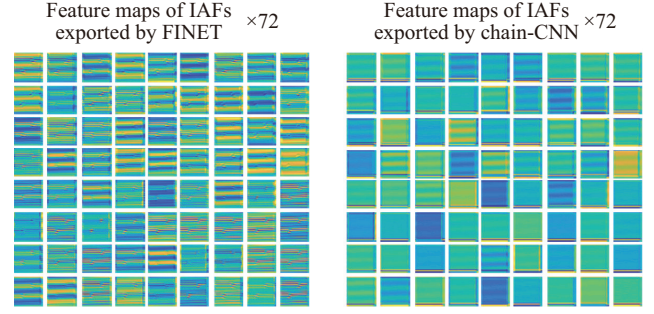


Fig. 12. The feature maps of FINET and chain-CNN.

TABLE VI
DETECTION ACCURACY OF PROPOSED METHOD IN COMPARISON WITH SOME STATE-OF-THE-ART WORKS

Comparative literature	Method	Detection accuracy(ϕ)		
		Software simulated	Measured	Mix
Sarlak [5]	MMG+perceptron	97.3%	NA	NA
Mishra [7]	ST+ANN	100%	NA	NA
Samantaray [9]	AEKF +PNN	96.8%	NA	NA
Mortazavi [13]	SWT+SVMs	NA	NA	97.2%
Lala [16]	EMD+ANN	over 98%	NA	NA
Proposed method	FINET+PSF	100%	100%	100%

unfamiliar sets. Therefore, FINET is used as the front-end, and the PSF method is proposed as the back-end detector structure. Software simulation and measured zero-sequence current waveforms (Table II) are input to the algorithm for identification. Detection results are shown in Table VI.

Table VI compares the detection accuracy of FINET+PSF methods for IAF and non-IAF zero-sequence current waveforms and some highly cited papers in the field of IAF detection. Different papers use different software simulation models, identification methods, and data sources. Thus, only final detection accuracy is compared, and detection accuracy (ϕ) expression of this paper is as shown in (5).

$$\phi = \frac{\text{Correct numbers of complete waveforms}}{\text{Total numbers of complete waveforms}} \times 100\% \quad (5)$$

The differences between the proposed method and other state-of-art methods are summarized as follows:

- 1) Other previous works only take short or long period waveforms in a complete fault waveform as identification samples. A fault index calculated from only a section of waveform may miss all the information that a complete waveform can provide. In this paper, the complete waveform is first divided into several detection windows and input to FINET to obtain a sequence of 0 or 1. The counting method of PSF is introduced to statistically analyse the sequence of 0 or 1 to improve the recognition accuracy of the complete waveform.
- 2) When there is no intermittent process in the IAFs, PSF can count consecutive 1 sequences to judge this is a long-term arc fault. The general method may ignore intermittent processes in IAFs, even including frequent faults and intermittent processes alternations. The PSF can produce statistics during the intermittent failure process. If the intermittent process is very short, the PSF method ignores

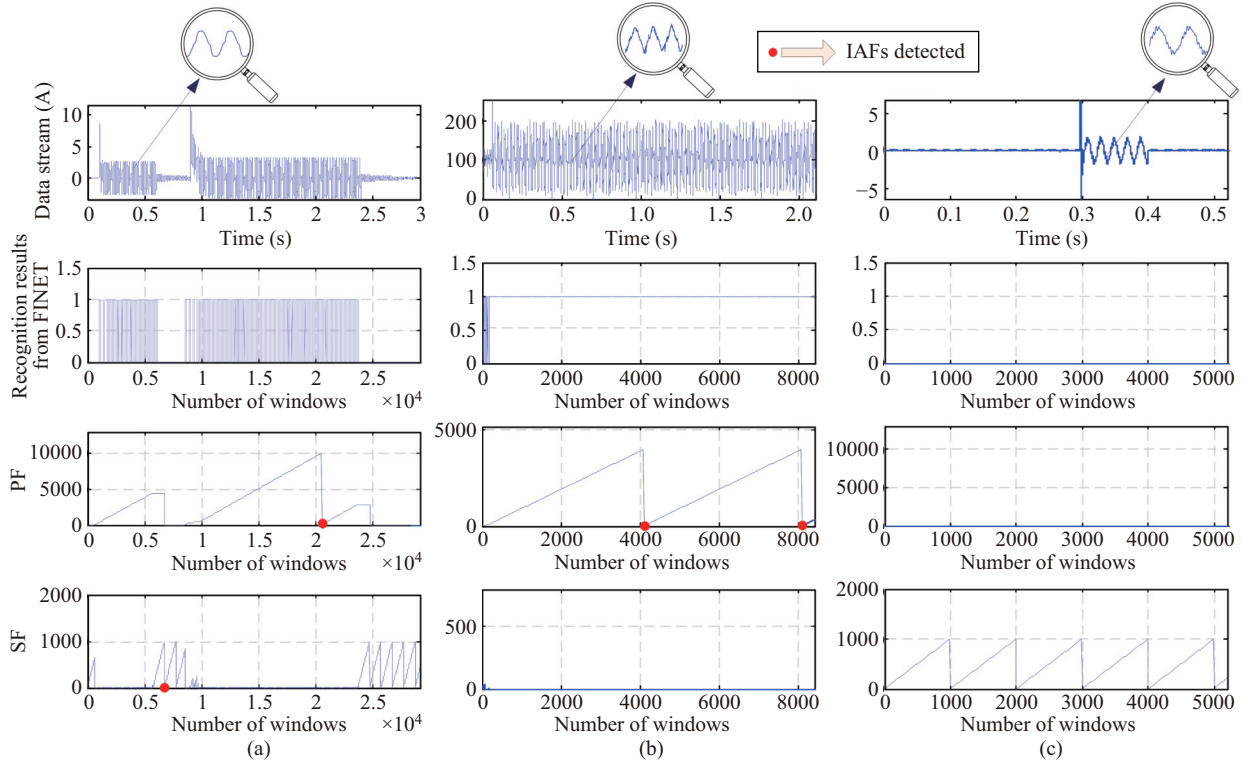


Fig. 13. Practice detection results curve. (a) Simulation model 3: IAFs. (b) Field data: IAFs. (c) Hardware experiment: single-phase-to-ground fault.

it and continues to count the subsequent 0 or 1 sequences; if the intermittent process is long, the PSF method clears and recounts after making a fault judgement in the previous statistical sequence. Compared with methods that do not divide the signal, this method provides improved immunity and eliminates signal redundancy.

Given the complex and variable on-site environment and fault conditions, it is difficult to select an appropriate threshold to determine fault occurrence. Different from traditional threshold triggering methods, the combination of FINET and PSF confers strong interference resistance and robustness. Thus, both simulation model data and measured data can be accurately identified. The method can provide real-time detection of abnormalities in the latest window data, and can confirm whether IAFs occur in a short detection period. Therefore, early warning of a serious ground fault can be implemented to enhance security of power distribution systems.

B. Engineering Practice and Implementation

The author's laboratory recording device is combined with a Jetson TX2 to implement a real-time detection system based on deep learning, as shown in Fig. 3(b). The device is applied to hardware experimental systems for testing, as shown in Fig. 13(c). For detecting the result of a single-phase ground fault in the hardware system, the first line is the zero-sequence current waveform of the single-phase ground fault, the second line is the 0 or 1 sequence output by FINET, the third line is the statistical curve of PF, and the fourth line is the statistical curve of SF. It can be seen from the 0 or 1 sequence that FINET can effectively distinguish between single-phase ground fault and

IAFs. Therefore, the statistical curve of PS does not change, and the statistical curve of SF keeps counting and clearing.

Similarly, Fig. 13(a) and Fig. 13(b) are respectively detection results of simulation and field complete waveforms. Taking Fig. 13(a) as an example, it can be known from the 0 or 1 sequence that FINET cannot judge all waveforms of all detection windows correctly, but through statistics of PF and SF, short-term IAFs are settled in the intermittent process, and long-term IAFs in the fault process are detected. Both short-term and long-term IAFs are accurately captured, and the early warning signal of the IAFs is outputted. The above experimental results indicate that FINET+PSF can minimize false negative rates of IAFs.

V. CONCLUSION

An accurate and reliable method for detecting intermittent arc faults is proposed based on combining the deep neural network FINET and a proposed PSF method. By extracting the real-time waveform window of zero-sequence current, IAFs and other events in the resonant grounding distribution system can be accurately distinguished without time-frequency decomposition. In the experiment, PSCAD/EMTDC was used to simulate three kinds of distribution network models to obtain software simulation data which is used for training and testing. In addition, hardware experimental data and field data are also tested to verify the proposed method's adaptability. The results show the proposed method has good adaptability since both of the simulation and measured data can be accurately detected. The proposed approach achieves IAF detection accuracy in excess of 90%, while that of FINET + PSF

is 100%, outperforming current state-of-the-art methods and indicating feasibility for practical engineering tasks.

REFERENCES

- [1] M. G. Adamiak, B. D. Russell, C. L. Benner, R. W. Dempsey, J. M. Parsick, and A. C. Depew, "Field experience with high-impedance fault detection relays," in *Proceedings of the 2005/2006 IEEE/PES Transmission and Distribution Conference and Exhibition*, 2006, pp. 868–873.
- [2] T. J. Yuan, F. Yin, D. Jian et al., *Single Phase Arc Grounding Over-voltage in Modern Distribution Network*, China Electric Power Press, 2017.
- [3] S. R. Ola, A. Saraswat, S. K. Goyal, S. K. Jhajharia, B. Rathore, and O. P. Mahela, "Wigner distribution function and alienation coefficient-based transmission line protection scheme," *IET Generation, Transmission & Distribution*, vol. 14, no. 10, pp. 1842–1853, May 2020.
- [4] A. Ghaderi, H. L. Ginn III, and H. A. Mohammadpour, "High impedance fault detection: A review," *Electric Power Systems Research*, vol. 143, pp. 376–388, Feb. 2017.
- [5] M. Sarlak and S. M. Shahrtash, "High impedance fault detection using combination of multi-layer perceptron neural networks based on multi-resolution morphological gradient features of current waveform," *IET Generation, Transmission & Distribution*, vol. 5, no. 5, pp. 588–595, May 2011.
- [6] W. C. Santos, F. V. Lopes, N. S. D. Brito, and B. A. Souza, "High-impedance fault identification on distribution networks," *IEEE Transactions on Power Delivery*, vol. 32, no. 1, pp. 23–32, Feb. 2017.
- [7] M. Mishra, P. Routray, and P. K. Rout, "A universal high impedance fault detection technique for distribution system using S-transform and pattern recognition," *Technology and Economics of Smart Grids and Sustainable Energy*, vol. 1, no. 1, pp. 9, Aug. 2016.
- [8] S. Gautam and S. M. Brahma, "Detection of high impedance fault in power distribution systems using mathematical morphology," *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 1226–1234, May 2013.
- [9] S. R. Samantaray, P. K. Dash, and S. K. Upadhyay, "Adaptive kalman filter and neural network based high impedance fault detection in power distribution networks," *International Journal of Electrical Power & Energy Systems*, vol. 31, no. 4, pp. 167–172, May 2009.
- [10] R. Bhandia, J. de Jesus Chavez, M. Cvetković, and P. Palensky, "High impedance fault detection using advanced distortion detection technique," *IEEE Transactions on Power Delivery*, vol. 35, no. 6, pp. 2598–2611, Dec. 2020.
- [11] M. Michalik, W. Rebizant, M. Lukowicz, S. J. Lee, and S. H. Kang, "High-impedance fault detection in distribution networks with use of wavelet-based algorithm," *IEEE Transactions on Power Delivery*, vol. 21, no. 4, pp. 1793–1802, Oct. 2006.
- [12] M. R. Haghifam, A. R. Sedighi, and O. P. Malik, "Development of a fuzzy inference system based on genetic algorithm for high-impedance fault detection," *IEEE Proceedings-Generation, Transmission and Distribution*, vol. 153, no. 3, pp. 359–367, May 2006.
- [13] S. H. Mortazavi, Z. Moravej, and S. M. Shahrtash, "A hybrid method for arcing faults detection in large distribution networks," *International Journal of Electrical Power & Energy Systems*, vol. 94, pp. 141–150, Jan. 2018.
- [14] H. Livani and C. Y. Evrenosoğlu, "A fault classification and localization method for three-terminal circuits using machine learning," *IEEE Transactions on Power Delivery*, vol. 28, no. 4, pp. 2282–2290, Oct. 2013.
- [15] I. Baqui, I. Zamora, J. Mazo'n, and G. Buigues, "High impedance fault detection methodology using wavelet transform and artificial neural networks," *Electric Power Systems Research*, vol. 81, no. 7, pp. 1325–1333, Jul. 2011.
- [16] R. Bhandia, J. de Jesus Chavez, M. Cvetković, and P. Palensky, "High impedance fault detection using advanced distortion detection technique," *IEEE Transactions on Power Delivery*, vol. 35, no. 6, pp. 2598–2611, Dec. 2020.
- [17] M. F. Guo, X. D. Zeng, D. Y. Chen, and N. C. Yang, "Deep-learning-based earth fault detection using continuous wavelet transform and convolutional neural network in resonant grounding distribution systems," *IEEE Sensors Journal*, vol. 18, no. 3, pp. 1291–1300, Feb. 2018.
- [18] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: A review and new perspectives," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 8, pp. 1798–1828, Aug. 2013.
- [19] I. V. Tetko, D. J. Livingstone, and A. I. Luik, "Neural network studies. 1. Comparison of overfitting and overtraining," *Journal of Chemical Information and Computer Sciences*, vol. 35, no. 5, pp. 826–833, Sep. 1995.
- [20] S. Hochreiter, Y. Bengio, P. Frasconi, and J. Schmidhuber, "Gradient flow in recurrent nets: the difficulty of learning long-term dependencies," in *A Field Guide to Dynamical Recurrent Neural Networks*, S. C. Kremer and J. F. Kolen, Eds. Piscataway: IEEE Press, 2001, pp. 237–244.
- [21] C. Szegedy, W. Liu, Y. Q. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 1–9.
- [22] K. M. He, X. Y. Zhang, S. Q. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 770–778.
- [23] Y. Xu, M. F. Guo, B. Chen, and G. J. Yang, "Modeling and simulation analysis of arc in distribution network," *Power System Protection and Control*, vol. 43, no. 7, pp. 57–64, Apr. 2015.
- [24] A. L. Lin, M. F. Guo, W. Gao, and J. Y. Huang, "Flexibility fault generating device of physical simulation system for distribution network," *Electrical Engineering*, no. 11, pp. 18–24, Dec. 2017.
- [25] H. F. Yang, C. Hong, Y. Gao, J. Y. Huang, B. Chen, and W. Gao, "Design and implementation of the control unit in a distribution network physical simulation system," *Power System Protection and Control*, vol. 45, no. 23, pp. 122–129, Dec. 2017.
- [26] Z. Y. Zheng, X. Y. Cai, W. Z. Wang, G. J. Yang, and B. H. Zhang, "Development of distributed feeder fault monitoring system in distribution network," *Electrical and Energy Management Technology*, no. 13, pp. 46–51, Sep. 2016.
- [27] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proceedings of the 25th International Conference on Neural Information Processing Systems*, 2012, pp. 1097–1105.
- [28] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 2261–2269.
- [29] V. Nair and G. E. Hinton, "Rectified linear units improve restricted boltzmann machines," in *Proceedings of the 27th International Conference on Machine Learning*, 2010, pp. 807–814.
- [30] C. M. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)*, New York: Springer, 2006.
- [31] J. Schmidhuber, F. Gers, and D. Eck, "Learning nonregular languages: A comparison of simple recurrent networks and LSTM," *Neural Computation*, vol. 14, no. 9, pp. 2039–2041, Sep. 2002.
- [32] K. Cho, B. van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2014, pp. 1724–1734.



Ye Tian obtained a Master's degree in Electrical Engineering from Yuan Ze University, Taiwan, China, in 2019, and a Master's degree in Electrical Engineering from Fuzhou University, China, in 2020. After graduation, he works in Fuzhou Power Supply Company of State Grid Fujian Electric Power Co., Ltd. His research interests include power distribution systems protection and application of machine learning in power distribution systems.



Mou-Fa Guo received the B.S. and M.S. degrees in Electrical Engineering from Fuzhou University, Fuzhou, China, in 1996 and 1999, respectively, and the Ph.D. degree in Electrical Engineering from Yuan Ze University, Taiwan, China, in 2018. Since 2000, he has been with Fuzhou University, where he is currently a Professor with the College of Electrical Engineering and Automation. His research interests include the information processing, protection control and flexible arc suppression of single-line-to-ground fault in distribution networks.



Duan-Yu Chen received a B.S. degree in Computer Science and Information Engineering from National Chaio-Tung University, Taiwan, China, in 1996, an M.S. degree in Computer Science from National Sun Yat-Sen University, Taiwan, China, in 1998, and a Ph.D. degree in Computer Science and Information Engineering from National Chiao-Tung University, Taiwan, China, in 2004. He was a Postdoctoral Research Fellow with Academia Sinica, Taipei, Taiwan, China, during 2004–2008. He was an Assistant Professor in the Department of Electrical Engineering, Yuan Ze University from 2009 to 2013. He is currently a Professor and Chairman of the Department of Electrical Engineering, Yuan Ze University, Taiwan, China. He received the Young Scholar Research Award presented by Yuan Ze University, Taiwan, China, 2012. His research interests include computer vision, video signal processing, content-based video indexing and retrieval, and multimedia information systems.