Received 3 November 2023; revised 23 February 2024 and 10 July 2024; accepted 30 September 2024.
Date of publication 10 October 2024; date of current version 21 October 2024.

Digital Object Identifier 10.1109/OAJPE.2024.3477630

Learning Power Systems Waveform Incipient Patterns Through Few-Shot Meta-Learning

LIXIAN SHI[1], QIUSHI CUI[1] (Member, IEEE), YANG WENG[2] (Senior Member, IEEE), YIGONG ZHANG[1] (Student Member, IEEE), SHILONG CHEN[3], JIAN LI[1] (Senior Member, IEEE), AND WENYUAN LI[1] (Life Fellow, IEEE)

[1]School of Electrical Engineering, Chongqing University, Chongqing 400044, China
[2]Department of Electrical, Computer and Energy Engineering, Arizona State University, Tempe, AZ 85281 USA
[3]Faculty of Electric Power Engineering, Kunming University of Science and Technology, Kunming 650500, China
CORRESPONDING AUTHOR: Q. CUI (qcui@cqu.edu.cn)

ABSTRACT Incipient faults (IFs) are abnormal states before the permanent failure of power equipment. IFs are typically transient and generally do not trigger the operation of relay protection devices. This leads the difficulty in capturing IF data from waveform monitoring or recording devices. However, traditional detection methods cannot achieve satisfactory performance when faced with limited data. Besides, some signal analysis methods based on waveform conversion to images cannot obtain understandable image data and cannot analyze both current and voltage signals simultaneously. To resolve these problems, a few-shot meta-learning framework for incipient fault detection (FSMLF-IFD) is proposed in this paper. For better data processing, a waveform image conversion strategy is proposed to convert waveforms into understandable images from the time domain perspective. Then, an adaptive image fusion strategy is developed to concurrently analyze voltage and current images. Next, at the meta-training stage, an adaptability-enhancing weighting initialization strategy is constructed to address the data differences between the meta-training stage and IF detection stage. Finally, an IF detection model based on convolutional neural networks (CNNs) is obtained through the fine-tuning process. In the numerical results, the IF detection and classification accuracy of FSMLF-IFD reached 0.9720 and 0.9840 based on simulation and field IF data, which validates the effectiveness of the proposed method.

INDEX TERMS Incipient fault detection, power quality, data scarcity, waveform image, few-shot learning, meta-learning.

NOMENCLATURE

$1-shot$	Use one sample for meta-training process.
$2-D$	Two-dimension.
CNN	Convolutional neural network.
DFA	Distribution fault anticipation.
DG	Distributed generator.
EFD	Early fault detection.
$FSMLF-IFD$	Few-shot meta-learning framework for incipient fault detection.
GASF	Gramian Angular Summation Field.
IF	Incipient fault.
LL	Line-to-line.
LLG	Line-to-line-to-ground.
LLLG	Three-line-to-ground.
LR	Logistic regression.
$Micro-PMU$	Micro-phasor measurement units.
PQ	Power quality.
RBF	Radial basis function.
RMS	Root mean square.
SLG	Single line-to-ground.
SNR	Signal-To-Noise Ratio.
SPM	space-phasor model.
ST	S-transform.
TSML	Task-sequencingmeta-learning.
WMU	Waveform measurement units.
WT	Wavelet transform.

I. INTRODUCTION

NOWADAYS, the integration of a significant proportion of renewable energy sources into power distribution and transmission systems poses challenges to grid stability and fault detection [1]. To ensure the safe operation of power systems, the development of intelligent fault detection or location strategies is imperative [2], [3], [4], [5]. Fault events in a power system can be divided into two types: unpredictable and predictable faults [6]. Generally speaking, unpredictable faults are unexpected events caused by bad weather, human misoperation, tree contacts, and other factors, which are usually random. On the contrary, predictable equipment failure events are generally not instantaneous. Some equipment faults such as cable joint failures could be predicted by IFs. The damage of these faults may not be fatal at first. As time goes by, equipment permanent failure events may eventually occur. Permanent failure may lead to a series of cascading faults, ultimately leading to widespread power outages [7].

IFs are precursory phenomenons before permanent failure of power equipment [8], [9]. Various categories of IFs exist, each exhibiting distinct features and failure patterns, often resulting in unique signatures. However, the challenge lies in the fact that some IFs may not trigger relay protection mechanisms and can self-clearing under specific circumstances, such as underground cables [10]. Furthermore, another challenge lies in the difficulty of distinguishing between IFs of certain devices and other faults. When an IF occurs, it is necessary to detect it, make appropriate decisions and activate appropriate alarm signals. For this reason, effective IF detection can help avoid catastrophic permanent failures of different devices, thus improving power supply reliability [8].

Due to the differences in the network structure, topology, equipment parameters, and other aspects of the distribution network with renewable energy sources, it is difficult to set unified IF detection rules. A feasible way of detecting IFs is to analyze the power quality (PQ) data [11], [12]. PQ monitors passively observe the dynamics of power systems [13]. In recent years, engineers and researchers in the field of power system protection and equipment testing have demonstrated that power equipment faults or failures, such as the arcing of cable joints and the unsuccessful synchronous closing control, can produce different signatures. Therefore, the analysis and mining of PQ data can help detect and identify IFs of power equipment.

In view of the above facts, many researchers have explored fault characteristics and detection methods of IFs. In general, characteristics of IFs are obtained by analyzing PQ waveforms from three perspectives: time domain, frequency domain [14], and time-frequency domain [15]. The magnitude and duration of fault waveforms are usually analyzed in the time domain. For instance, Kulkarni et al. [16] proposes an IF location algorithm of underground cables that considers the time domain analysis and fault arc voltages. Besides, some signal decomposition methods based on frequency domain and time-frequency domain, such as Fourier transform [14], wavelet analysis [17] and S-transform (ST) [6], [18], are applied to fault feature extraction. To effectively mine extracted fault features, machine learning methods have recently been widely applied in IF detection [19], [20], [21]. A machine-understandable pre-training model based on fast Fourier transform and natural language processing technique is established for general-purpose IF detection in [14]. Considering ST and Kullback-Leibler divergence, [6] adopts two feature measures to obtain seven peculiar features from voltage waveforms of abnormal phases recorded by PQ meters at substations. A two-step strategy for IFs monitoring of underground cables has been proposed based on cumulative sum and adaptive linear neuron [10]. However, the previously mentioned approaches focus on analyzing the discrete fault waveforms, which only extract fault characteristics from one-dimensional waveform series data.

With the development of artificial intelligence, some excellent image recognition technologies have emerged like mushrooms after rain. Converting waveform signals into images and using these advanced technologies can improve the accuracy of fault identification [20], [22], [23], [24], [25], [26]. Salles and Ribeiro [27] transform voltage signals into 2-D images with time-frequency representation by continuous wavelet transform (WT). Similar image conversion methods can also be found in [24] and [25]. Izadi and Mohsenian-Rad [20] and [26] have converted the synchro-waveform data into synchronized Lissajous images to mine the power event characteristics from the time domain. Besides, the space-phasor model (SPM) and Gramian Angular Summation Field (GASF) are applied to convert waveforms into images in literature [23] and [28], respectively. However, the above conversion methods cannot obtain understandable image data focused on fault waveform distortion. Moreover, both current and voltage images cannot be analyzed simultaneously.

In addition, since most traditional detection methods are supervised learning algorithms, these methods usually need a lot of fault data with labels [29]. A hybrid algorithm based on decision trees and computational intelligence is trained through 162 samples to diagnose IFs of power transforms in Menezes et al. [30]. An intelligent learning method based on probability uses one hundred and twenty abnormal event data to realize the classification of IFs in the literature [19]. Andresen et al. [31] discuss that many machine learning methods are trained using the collected long-term PQ data to predict the fault in the grid. However, the reality is that data are scarce on actual incipient failures. This dilemma reveals that traditional supervised learning methods are not effective in detecting IF under the limited fault sample.

The aforementioned dilemma encourages researchers to explore the new perspective of fault data mining and to construct an IF detection model that can learn IF fault characteristics from limited samples. Xiong et al. [19] propose a human-level concept learning framework based on Bayesian

TABLE 1. The contribution compared to existing literature.

Literature	Method	Limitation	Contribution
[20], [22]–[26]	Image generation based on various methods	Hard to understand and no specific meaning	Waveform image conversion strategy
		Only analyzing current or voltage	Adaptive image fusion strategy
[32], [33]	Meta-learning framework	Neglecting the data differences	Adaptability-enhancing weighting initialization strategy

learning to detect and classify subcycle and multicycle IFs under data scarcity. However, this framework requires strong prior knowledge. Furthermore, some researchers have focused on the meta-learning framework for fault detection [32], [33]. Hu et al. [32] propose a task-sequencing meta-learning method to address the few-shot fault diagnosis problem. However, these studies neglect the differences between meta-training and realistic fault detection tasks.

To address the above issues, a few-shot meta-learning framework for incipient fault detection (FSMLF-IFD) is constructed. This method includes three contributions: a waveform image conversion strategy, an adaptive image fusion method, and an adaptability-enhancing weighting initialization strategy. To clearly compare our contributions with previous research, a taxonomy table is provided in Table 1. The detailed contributions are summarized as follows:

- A PQ waveform image conversion strategy is proposed. This strategy converts PQ waveforms into understandable waveform images from the time domain perspective.
- An adaptive image fusion strategy based on the amplitude deviation of PQ waveforms is developed. This strategy can fuse current and voltage waveform images into one image to simultaneously process the current and voltage data.
- A few-shot meta-learning framework for incipient fault detection (FSMLF-IFD) is constructed. This method is based on the meta-learning framework to learn the fault characteristics of IFs with limited samples. In this method, an adaptability-enhancing weighting initialization strategy is proposed. This strategy can address the data differences between the meta-training stage and IF detection stage. This enhances adaptability and robustness during the meta-training stage.

The rest of the paper is organized as follows. Some IF characteristics are analyzed in Section II. The data preparation process of FSMLF-IFD is provided in Section III. The proposed FSMLF-IFD is demonstrated in Section IV. The numerical results are displayed in Section V. Finally, the conclusion is present in Section VI.

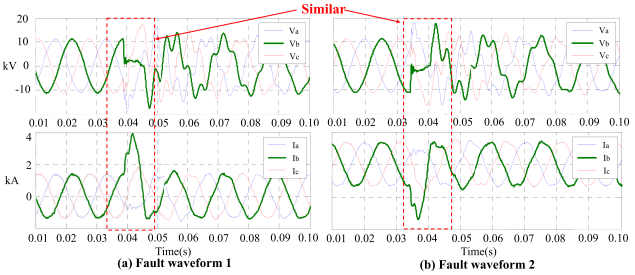


FIGURE 1. Typical IF waveforms of cables in a distribution network [34].

II. ANALYSIS OF INCIPIENT FAULTS CHARACTERISTICS

IFs refer to transient faults that occur before the complete failure of the equipment. If IFs are not dealt with, they may develop into permanent failures. According to [9], the fault characteristics of IFs can be captured from PQ waveforms. Fig. 1 presents two current and voltage waveforms of cable termination failures [34]. As we can see, although the inception time of fault waveform 1 and fault waveform 2 is different, they present similar waveform variation trends. This implies that the detection and identification of IFs can be achieved through the analysis of voltage and current waveforms. Consequently, the development of a machine-learning algorithm for IF detection using voltage and current data is a viable approach. In addition, due to the brief duration of faults and the low level of fault current, the protective device is typically not activated. As a result, capturing the IF waveforms becomes challenging. Consequently, the machine-learning model for IF detection can not only integrate IF characteristics but also solve the limited sample problem. In addition, this paper focuses on single-phase IFs. The IF of cables is caused by insulation damage and degradation, which generally appears in a single phase. The probability of simultaneous failure in all three phases is low. The related theories of IFs are usually analyzed based on single-phase faults [17], [35].

III. DATA PREPARATION PROCESS OF FSMLF-IFD

In this section, the data preparation process of FSMLF-IFD is presented and the remaining two parts will be introduced in the next section. Firstly, the waveform image conversion strategy is proposed to convert discrete waveform data from the time domain into understandable images. This conversion strategy can mine time-varying characteristics of amplitude and deformation characteristics of waveforms. Furthermore, this strategy is more sensitive to changes in the amplitude of waveform signals. After converting high-amplitude waveforms of IFs into images, the IF image data becomes easier to distinguish compared to discrete waveform data. Secondly, the voltage and current images are fused into a image through

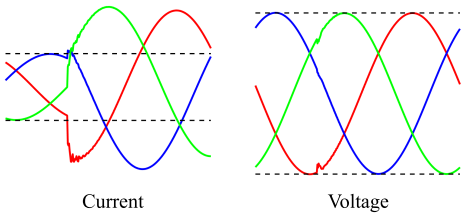


FIGURE 2. A standard waveform image after conversion. Simulated event: load adding.

the adaptive image fusion strategy. Finally, based on these two strategies, a fault database is constructed.

A. WAVEFORM IMAGE CONVERSION

To transform waveform data into comprehensible waveform images a waveform image conversion strategy is proposed. Firstly, a sliding window is employed to capture the current and voltage waveforms. It is worth noting that 3-phase current and voltage waveforms are considered at the same time for detecting equipment failures. Because the amplitude information is very important for fault detection, the amplitude information of voltage or current signal should be reflected in generated waveform images. Secondly, two dotted lines are added to the waveform image to indicate the amplitude value under the steady-state condition. Thirdly, utilizing the predetermined image dimensions, graphical representations of the current and voltage waveforms can be produced.

Based on the above three steps, discrete fault or abnormal waveform data of a sliding window can be transformed into waveform images. For instance, the current and voltage images of load switching by circuit breaker operation are shown in Fig. 2. The red, black, and blue curves represent the instantaneous values of 3-phase voltage or current, while the black dashed line illustrates the magnitude value during stable operating conditions.

Compared to raw waveform data, the benefits of waveform images are as follows.

- Transforming into waveform images for analysis can avoid the cumbersome process and difficulty in determining parameters associated with signal processing methods [22]. Analyzing raw waveform data typically relies on signal analysis techniques such as waveform transform (WT), ST, etc. These methods not only require manual parameter settings but also involve a complex analysis process. Improper parameter settings often lead to poor classification accuracy. In contrast, using waveform images for analysis eliminates parameter setting issues and simplifies the process.
- Intuitiveness. Converting waveforms can visually depict essential information about waveforms such as amplitudes and distortion. Fig. 3 shows current waveform images of IFs. As depicted in Fig. 3, our image-conversion method effectively captures waveform amplitudes and distortions.

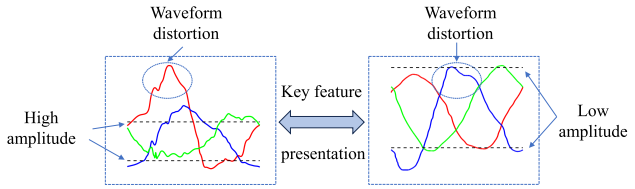


FIGURE 3. Two current waveform images of IFs.

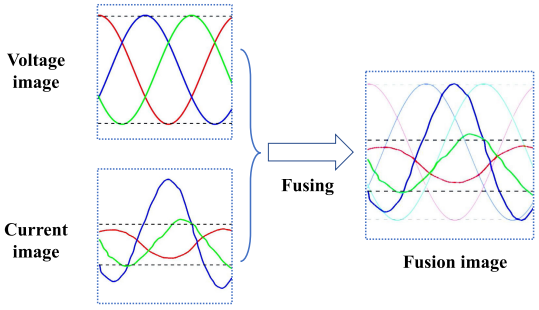


FIGURE 4. Fusion process of current and voltage waveform images of the cable IF.

B. ADAPTIVE IMAGE FUSION STRATEGY

Since the characteristics of faults are reflected in both voltage and current signals [8], it is advantageous to concurrently analyze voltage and current data. Consequently, we consider merging the voltage and current waveform images into one image as the input of the detection model. For a certain fault type, the fusion process of voltage and current images is shown in Fig. 4. The current and voltage waveforms are fused into one image based on the proposed image fusion strategy, which cannot be achieved using raw data. The fused images alleviate the challenge of separately analyzing current and voltage data. Then, the detailed fusion process is described as follows.

It is reasonable that prominent fault characteristics are displayed in significant waveform signal deformation. For the waveform images, the clearer the trend of waveform changes in the image, the larger the pixel value. The clear waveform variation can be more effectively learned by subsequent models. Therefore, the purpose of the adaptive image fusion strategy is to ensure that the distorted current or voltage waveform is more prominently represented in the fused waveform image.

In this paper, the difference in root mean square (RMS) value between the fault voltage or current waveform and the steady-state waveform is used to describe the amplitude deviation. Assuming that the amplitude deviation of voltage is represented by V_{AD}, its calculation formula is as follows:

$$V_{AD} = |V_{RMS} - V_{nom}|, \tag{1}$$

where V_{RMS} is the RMS value of the voltage waveforms, and V_{nom} represent RMS values of the steady-state voltage waveform. $|\cdot|$ denotes the absolute value. To ensure λ_V and λ_I are greater than 0, the absolute value should be applied to

$V_{RMS} - V_{nom}$. For a sampling window, the voltage or current series can be considered as the sum of time-varying sinusoids and noise, as presented in Eq. (2).

$$V_n = v_m \sin(\omega_n + \phi_n) + d_n, \tag{2}$$

where v_m is the max magnitude. ω_n and ϕ_n are the angular frequency and phase angle of the nth sampling instant, respectively. The noise d_n represents noise from the measuring device. Under normal circumstances, ω_n is the fundamental component. The calculation formula of V_{RMS} is as follows:

$$V_{RMS} = \sqrt{\frac{1}{k+1}\sum_{i=0}^{k} V_i^2}, \tag{3}$$

where k is the number of sampling points within a sampling period. The amplitude deviation of current I_{AD} can be calculated through the above process.

Then, given the voltage image x_V and current image x_I of a certain fault, the adaptive fusion strategy is realized as follows:

$$\hat{x} = \lambda_V x_V + \lambda_I x_I, \tag{4}$$

where $\hat{x}$ represents the combined waveform image data. x_V and x_I represent the voltage and current images of a specific fault type, respectively. λ_V and λ_I represent the fusion weights of voltage and current images, respectively. It is worth noting that the scaling values of the two images represent fusion weights. Then, the detailed calculations of fusion weights according to the amplitude deviation are as follows:

$$\lambda_V = \frac{V_{AD}}{V_{nom}}, \tag{5}$$

$$\lambda_I = \frac{I_{AD}}{I_{nom}}, \tag{6}$$

where V_{AD} and I_{AD} are the amplitude deviation of voltage and current waveforms.

C. CONSTRUCTION OF THE FAULT DATABASE

In the meta-training phase of the proposed FSMLF-IFD, the initialization parameters of the IF detection model are obtained by training numerous classification tasks. This process also necessitates a substantial amount of fault data. Then, the parameters are fine-tuned to learn the characteristics of IFs. The data distribution of meta-training and meta-testing is preferably similar. For the task of IF detection, the classification tasks during the meta-training stage should be as relevant as possible to the fault classification of power systems. Therefore, we need to build a fault database that includes various types of fault and event data. This will be based on waveform image conversion and adaptive image fusion strategies for the meta-training stage of FSMLF-IFD.

In the fault database, both fault and non-fault waveform image data are included. Fault events include single line-to-ground (SLG) faults, line-to-line-to-ground (LLG) faults, three-line-to-ground (LLLG) faults, and line-to-line (LL) faults. Non-fault events include steady-state, increasing load, and removing load. To ensure the diversity of the dataset, the inception angle of each fault event is considered in six cases: 0°, 60°, 90°, 120°, 180°, and 270°. The inception angle is calculated with respect to the voltage of Phase A. To simulate the aforementioned events, a power system distribution network is constructed, and detailed specifics are outlined in Section V. Through the simulation model, a significant amount of fault and non-fault discrete waveform data is acquired. Then, the simulation discrete current and voltage data of various events are converted into waveform images. Consequently, a fault database containing a large amount of fault information is established. By mining the fault database during the meta-training stage of FSMLF-IFD, a set of initialization parameters for the IF detection model is obtained. The parameters contain rich fault information and can be easily fine-tuned for a new IF detection task.

IV. PROPOSED IF DETECTION METHOD BASED ON FEW-SHOT LEARNING FRAMEWORK

In this paper, we propose an intelligent strategy to detect and identify IFs for the stability of the power system. Considering the scarcity of IF samples, we aim to train an IF detection model that can learn from a small number of IF samples and capture the characteristics of failure equipment. Therefore, this paper proposes FSMLF-IFD.

The schematic diagram of the proposed IF detection framework is shown in Fig. 5. The framework is divided into three parts: data preparation, meta-training, and IF detection. Firstly, for data preparation, the fault image database based on the waveform image conversion and the adaptive image fusion strategies is constructed. This fault database is used for the meta-training process. Secondly, many classification learning tasks are constructed and trained during the meta-training stage. This stage aims to obtain excellent initialization parameters for neural networks. An adaptability-enhancing weighting initialization strategy is proposed to ensure that the initialization parameters are suitable for IF detection. Finally, at the IF detection stage, the detection model with initialization parameters is fine-tuned through a small number of IF images. Through the fine-tuning process, an IF detection and identification model is obtained. Then, this detection model is integrated into the PQ monitoring system. IFs and other disturbances can be detected and identified. Since the data preparation process is introduced in Section III, the subsequent sections provide detailed descriptions of the remaining components of FSMLF-IFD.

A. META-TRAINING STAGE AND ADAPTABILITY-ENHANCING WEIGHTING INITIALIZATION STRATEGY

The goal of the few-shot learning framework is to obtain good initialization parameters for an IF learner based on the fault database. During the meta-training stage of FSMLF-IFD, the learner with initialization parameters has strong adaptability and good generalization performance on new fault types.

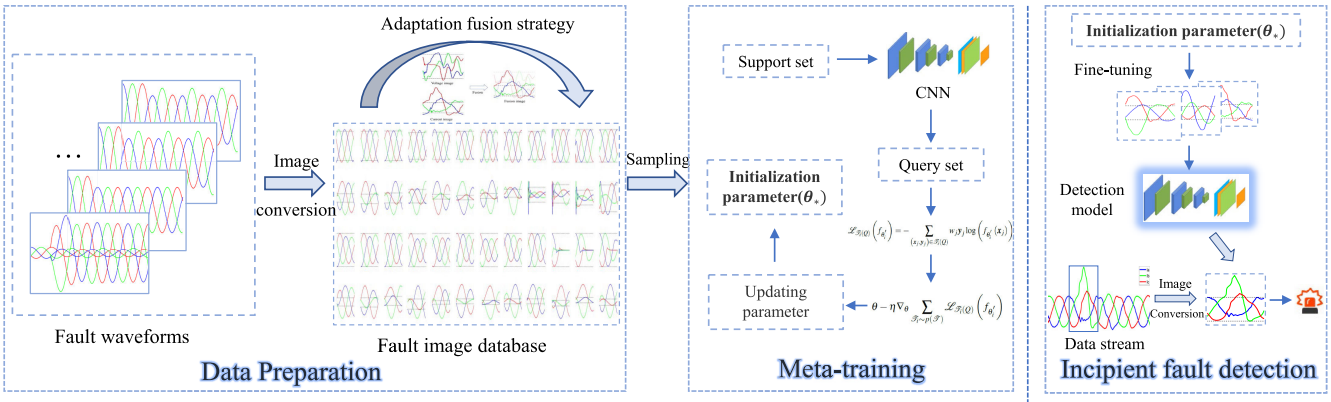


FIGURE 5. **Schematic diagram of the proposed IF patterns learning approach.**

When used for detecting IFs, the learner becomes the detection model after fine-tuning with a few IF waveform images.

During meta-training, the learner is trained on a set of tasks sampled from the constructed fault database. In other words, various fault classification tasks are considered as the training set at this stage, which is significantly different from the traditional machine learning setting. Convolutional neural networks (CNNs) are chosen as the learner for FSMLF-IFD. The reason is presented as follows. On the one hand, due to CNN's operation principle of extracting image features by continuously moving convolutional kernels forward, it excels in handling generated waveform image data. CNNs leverage the forward movement of convolutional kernels to effectively capture and process the temporal correlations of waveform images, thereby enhancing the accuracy of early fault detection. On the other hand, convolutional and pooling layers of CNNs can automatically learn local patterns and global features of waveform images without the need for manual feature extractor design. The distortion of IFs can be adaptively learned.

The dataset at the meta-training stage consists of two parts: the support set S and the query set Q. The S is utilized to train the parameters of the learner, while Q is employed to update the parameters of the learner. In the training process based on S, the loss function is the same as that in reference [36]. During the parameter updating process, we propose an adaptability-enhancing weighting initialization strategy. This strategy integrates the data characteristics of IFs into the meta-training process, focusing on obtaining initialization parameters for IF detection.

The specific details of the training process can be found in [36]. Then, the updating process and an adaptability-enhancing weighting initialization strategy will be described. Suppose θ_j' is the obtained parameters of the support set S, the loss $\mathcal{L}_{\mathcal{T}_j(Q)}\left(f_{\theta_j'}\right)$ is calculated through the corresponding query sets Q. The calculation process is based on the cross-entropy loss and gradient descent. When testing tasks differ from training tasks, it is necessary to employ a specific strategy to enhance the initialization parameters. During meta-training, the diversity of tasks should be ensured to achieve improved few-shot learning and generalization capabilities. For this reason, the image database for training initialization parameters contains numerous other permanent fault images. The fault current of these faults varies significantly, leading to the activation of relay protection devices. However, in reality, the fault currents of the distribution network with renewable sources are not significant. Due to the significant difference in fault amplitudes, the fault images exhibit remarkable variations in feature space. Therefore, different tasks have varying degrees of importance for the parameter update process. Tasks similar to minor fault currents of IFs should be prioritized, and the significance of permanent failure should be given minimal weight. Therefore, an adaptability-enhancing weighting initialization strategy for IFs is proposed.

The RMS value of the current is used to assess the degree of magnitude variation. A larger RMS value indicates a greater change in amplitude. The weight of the ith classification task loss is calculated as follows:

$$w_i = \frac{\sum_{n=1}^{N} e^{\Delta RMS_n}}{e^{\Delta RMS_i}}, \tag{7}$$

where $\Delta RMS = I_{RMS} - I_{nom}$ and N is the number of classication. To this end, based on the weighting initialization strategy, the loss $\mathcal{L}_{\mathcal{T}_j(Q)}\left(f_{\theta_j'}\right)$ in the query sets Q is calculated as:

$$\mathcal{L}_{\mathcal{T}_j(Q)}\left(f_{\theta_j'}\right) = -\sum_{(\boldsymbol{x}_i,\boldsymbol{y}_i)\in\mathcal{T}_j(Q)} w_i \mathbf{y}_i \log\left(f_{\theta_j'}(\boldsymbol{x}_i)\right). \tag{8}$$

The learner parameters are updated via stochastic gradient descent as follows:

$$\boldsymbol{\theta}_* \leftarrow \theta - \eta\nabla_\theta \sum_{\mathcal{T}_j\sim p(\mathcal{T})} \mathcal{L}_{\mathcal{T}_j(Q)}\left(f_{\theta_j'}\right), \tag{9}$$

where θ_* is the initialization parameter for IF detection, η is the meta step number of the updating process, and $\leftarrow$ indicates that the updated value of θ is assigned to θ_*. Until a superior initialization parameter suitable for IF detection is achieved, the aforementioned procedure is repeated.

We incorporate IF data properties into the meta-learning approach, making it more suitable for IF detection tasks.

B. IFs DETECTION BASED ON OBTAINED INITIALIZATION PARAMETERS

To achieve IF detection, an IF detection model is necessary. The IF detection model can be obtained by fine-tuning the initialization parameters θ_* with the CNN model. The CNN model with θ_* has fast adaptability and is easily fine-tuned to learn the IF characteristics. The data, parameters, optimizer, and epoch of the fine-tuning process are described in detail. The data of the fine-tuning process are a few waveform images of the multiple categories, including IF equipment and other types (e.g., steady state, disturbance, etc.). These data construct a classification task and are used for fine-tuning the CNN model with the initialization parameter θ_*. For parameters of the fine-tuning process, the convolutional, pooling, and final classification layers of CNN are used for fine-tuning. The optimizer of the fine-tuning process is Adam. The epoch of the fine-tuning process is 100. Through the fine-tuning process, the CNN model becomes the IF detection model, capable of distinguishing the IFs, steady state, and disturbances. It should be emphasized that if the data of multiple fault devices are used to fine-tune initialization parameters θ_*, the obtained detection model can detect multiple IFs at the same time. Then, we can integrate this IF detection model into the event-triggered analysis (e.g. fault recorders, PQ analyzers, etc.) and gap-less monitoring devices (e.g. waveform measurement units (WMU), micro-phasor measurement units (micro-PMU), etc.), to achieve IF detection and identification.

For the detection process, a sliding window with a period of length N is utilized to capture voltage and current waveform data for analysis. In practical scenarios, the initiation mechanisms vary across different application scenarios. When applied to the relevant event-triggered devices, the proposed method involves analyzing recorded waveforms by the continuously sliding window to thoroughly inspect all the voltage and current waveform data. In the case of gap-less monitoring equipment, an appropriate triggering mechanism needs to be defined. For instance, using voltage data as a reference, the sliding window will be activated when the supply voltage drops or increases to a certain value, such as 0.8 pu or 1.1 pu. Then, waveform image conversion and adaptive image fusion strategies are implemented to get a waveform image. Next, the obtained waveform image is fed into the IF detection model, and the monitoring result is obtained. When an IF occurs multiple times in the distribution system, the alarm signal will be transmitted to the system operator instead of triggering relay systems. Because typical IFs are sub-cycle or multi-cycle, the above detection process is also suitable for sub-cycle faults. For multi-cycle IFs, we consider dividing the multi-cycle IF into multiple sub-cycle faults. If a fault occurs in multiple consecutive cycles, this fault is determined as a multi-cycle fault. Based on this idea, even if the fault is not detected in the first cycle, it can be detected by subsequent cycles.

V. NUMERICAL RESULTS

In this section, the proposed FSMLF-IFD is evaluated through experiments in various aspects. We use realistic IF data of cables and simulation systems to verify the effectiveness of the proposed method. The realistic data are collected from power quality monitoring from [9] and [19]. The sampling rate of collected data is 6.4 kHz. The length of the sliding window is set as 10 cycles.

Several systems have been used for testing our method. A distribution network system shown in Fig. 6 [37] is the benchmark model. All testing distribution network systems are simulated in MATLAB/Simulink. The voltage class and frequency of this system are 25 kV and 60 Hz, respectively. Two distributed generators (DGs) are included in this benchmark model. The types of DGs A and B are synchronous generators and wind farms, respectively. The synchronous generator is rated at 4.17 kV, 9 MVA. The wind farm is rated at 575 V, 6.6 MVA, and is based on the type IV wind turbine model. Based on this distribution model, we obtain some simulation IF data. In detail, some settings of simulation models are as follows:

- Fault type and location. The simulation fault types of IF are divided into sub-cycle and multi-cycle. The fault locations are set to F1, F2, and F3 (red mark in Fig. 6). The measurement points are set as the upstream buses of fault locations at nodes B-2, B-8 and B-21. The impedance between three faults and meter points is $0.4757+j1.618$, $0.1815+j0.6712$, and $0.6222+j2.1160$, respectively.
- IF simulation model. In this paper, we employ Kizilcay's arc model [38] to simulate the IFs of cables. The equation of this arc model is presented as:

$$\frac{dg\,(t)}{dt} = \frac{1}{\tau}\left(\frac{\left|i_f\,(t)\right|}{u_0 + r_0\left|i_f\,(t)\right|} - g\,(t)\right) \tag{10}$$

$$g\,(t) = \frac{u_f\,(t)}{i_f\,(t)} \tag{11}$$

  where $g\,(t)$ is the arc conductance, $i_f\,(t)$ is the arc current, $u_f\,(t)$ is the arc voltage, τ is the arc time constant, r_0 is the characteristic arc resistance, u_0 is the characteristic arc voltage. The range of these parameters are set as [19]: $\tau = 0.2 \sim 0.4$ ms, $u_0 = 300 \sim 400$ V, $r_0 = 0.01 \sim 0.015\Omega$. The three parameters were independently set each time to values within their respective ranges with an equal probability.
- Noise addition. We add the random Gaussian noise into the simulated current and voltage waveform to simulate load changes, communication disturbances, etc. The Signal-To-Noise Ratio (SNR) is randomly set as 20, 30, and 40 dB.

Because of page limit, only some experimental results are based on simulation IF data. In the paper, the IF detection

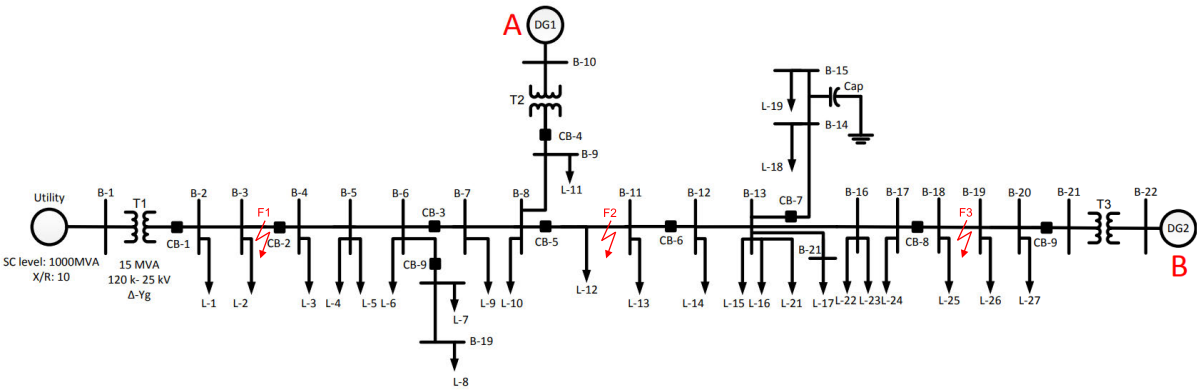


FIGURE 6. **Single line diagram of used distribution network model [37].**

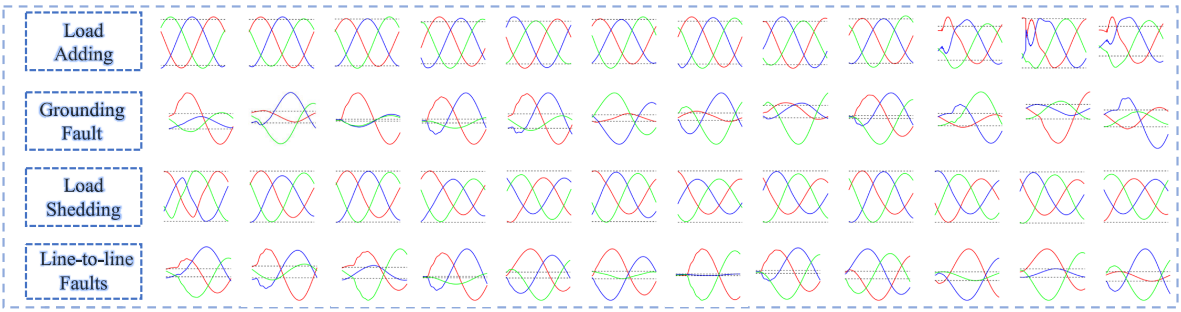


FIGURE 7. **Visualization of the fault current image before the image fusion. The database contains four types of events: load adding, grounding fault, load shedding, and line-to-line faults. Load adding and load shedding are realized by implementing circuit breakers. Grounding faults include single line-to-ground (SLG) faults, line-to-line-to-ground (LLG) faults, and three-line-to-ground (LLLG) faults. It is worth noting that the constructed fault database is used for training the initialization parameters.**

performance is evaluated using Accuracy, Precision, Recall rate, and F1 score. Furthermore, to obtain well-performing detection models, some critical parameters need to be set in advance. To avoid overfitting and ensure optimal parameters, we used a small-scale validation set to determine hyperparameters. This validation set is from the constructed fault database. The experimental platform of this paper is NVIDIA GeForce RTX 1650. All codes are implemented based on Pytorch. The size of waveform image is 256×256.

A. FAULT DATABASE CONSTRUCTION BASED ON A DISTRIBUTION NETWORK

To implement the meta-training stage of FSMLF-IFD, it is essential to construct a fault database that includes various types of faults and events. This fault database is used for the meta-training stage. In this subsection, the construction process of the fault database based on the distribution network model is described in detail.

As mentioned in the previous section, various fault and non-fault events are simulated using a distribution network model to obtain fault waveform image data. Several buses are considered for detecting current and voltage waveform data. For example, we can set the output at buses 2 and 14 to extract two different output behaviors under the same fault and event configurations. This way, we can analyze the impact of faults and events on different buses. In our research, we typically position the measurement points at the buses to the generators and the buses from the fault and event location. This allows us to observe changes near the generators and at more distant locations within the grid. Based on the above description, some current waveform images of the fault database are shown in Fig. 7. For a specific fault or event, the waveform image in the database displays some variations. The reason is that the fault and event waveform images are from different measurement buses, phases, or locations.

B. PERFORMANCE UNDER THE TRANSITION TO WAVEFORM IMAGE PROCESSING

In general, traditional methods analyze discrete current and voltage signals to locate or detect IFs of power equipment. However, this strategy cannot capture the dynamic time-varying trend of the waveform as a whole. For this purpose, discrete waveform data is converted into waveform images. This subsection will discuss the advantages of converting discrete waveform data into waveform images. Table 2 shows the classification performance comparison of

TABLE 2. Comparison between time-frequency methods and image processing methods.

Model	Accuracy	Precision	Recall	F1
WT+RBF [39]	0.6850	0.7765	0.7023	0.7612
Hybrid-SVM [6]	0.7258	0.8333	0.7465	0.7875
Waveform-CNN	0.6850	0.7222	0.6909	0.7060
Waveform image-LR	0.6060	0.8194	0.6060	0.6968
Waveform image-SVM	0.6969	0.7934	0.6969	0.7428
Waveform image-CNN	0.7576	0.8596	0.7743	0.8147

simulation data between the CNN model and two traditional fault detection strategies. The inputs of the CNN model are waveform images obtained from waveform conversion and adaptive image fusion strategies. In [39], WT and radial basis function (RBF) networks are combined to analyze the power quality disturbances. In addition, a hybrid method based on ST and support vector machines (SVM) is proposed for IF detection in [6]. We set a three-class classification problem to evaluate the classification performance of detection models. The three classes are normal, disturbance (non-IFs), and IFs of cable.

As shown in Table 2, two detection strategies based on signal processing cannot achieve good results and the classification accuracy is only 0.6850 and 0.7258. Waveform-CNN represents the use of one-dimensional CNN to learn raw waveform data directly, but its performance is not as good as the other two signal processing methods. Compared with these results, the accuracy of Waveform image-CNN is improved by 10.58%, 4.38%, and 10.58%, which shows the superiority of the waveform conversion and image fusion strategies. Waveform image-LR and Waveform image-SVM represent the identification of generated waveform images using LR and SVM, respectively. Compared with CNN, two machine learning methods cannot obtain good identification performance. In terms of accuracy, the two machine learning methods LR and SVM are 20.01% and 8.01% lower than CNN. This comparison results demonstrate the necessity of CNN for IF detection and identification. Furthermore, the results in the table are based on a small number of samples, which is the reason for poor performance.

C. STRONG LEARNING ABILITY AND ADAPTABILITY OF FSMLF-IFD

We use a three-class classification problem to evaluate learning ability and adaptability of FSMLF-IFD. Table 3 shows the multi-classification performance of different models for real and simulation IF data. Hybrid-SVM [6] and TSML [32] are two state-of-the-art detection methods. Compared with FSMLF-IFD, the hybrid approach [6] without considering limited samples cannot achieve good performance. Although [32] achieves decent performance, it is still not as excellent as our method. In terms of real data,

TABLE 3. IF identification performance comparison of different models for real and simulation IF data.

Model	Data	Accuracy	Precision	Recall	F1
Hybrid-SVM [6]	Real	0.7258	0.8333	0.7465	0.7875
	Sim	0.7058	0.8452	0.7066	0.7697
ResNet-18	Real	0.8788	0.9111	0.8825	0.8966
	Sim	0.8182	0.8823	0.8035	0.8410
CNN	Real	0.7879	0.8703	0.7654	0.8145
	Sim	0.7576	0.8596	0.7743	0.8147
TSML [32]	Real	0.9350	0.9406	0.9488	0.9447
	Sim	0.9490	0.9528	0.9501	0.9514
FSMLF-IFD	Real	0.9720	0.9760	0.9843	0.9801
	Sim	0.9840	0.9870	0.9887	0.9878

Note: The results are the average values of multiple experiments.

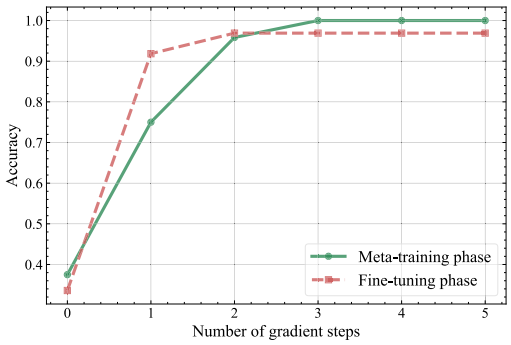


FIGURE 8. Accuracy variation of training and testing phase during gradient updating.

the accuracy is 3.96% lower than FSMLF-IFD. The image conversion and adaptability-enhancing weighting initialization strategy may be the reason for performance differences, which verifies the advantages of FSMLF-IFD. In addition, the classification accuracy of CNN is only 0.7879 in Table 3. With the help of FSMLF-IFD, satisfactory performance has been achieved, and its improvement in Accuracy, Precision, Recall rate, and F1 score are 23.37%, 8.70%, 19.30%, and 24.80%, respectively. This improvement strongly proves the learning ability of FSMLF-IFD.

To demonstrate the effectiveness of FSMLF-IFD in rapidly learning new samples, Fig. 8 illustrates the accuracy variation process of the proposed method with gradient updating during the meta-training stage and IF detection stage of an epoch. The data learned by the model during the meta-training process is the fault database, and the IF detection stage involves a fine-tuning process using the IF images. As shown in Fig. 8, the blue solid line represents the accuracy variation of the meta-training. The accuracy tends to be stable and increases from 0.375 to 1 with only 3 gradient steps. It indicates that the model has a strong adaptive learning ability. Additionally, the red dotted line represents the change in accuracy resulting

TABLE 4. The effectiveness of FSMLF-IFD for bidirectional power flow.

Situation	Accuracy	Precision	Recall	F1
Case 1 (forward)	0.9854	0.9890	0.9854	0.9860
Case 2 (reverse)	0.9690	0.9775	0.9690	0.9739

from fine-tuning in the few-shot learning. Similarly, the test accuracy increases significantly during the fine-tuning process with only a few gradient steps. This not only demonstrates the strong and rapid adaptation of the model but also underscores the effectiveness of the fault database.

To verify the effectiveness of FSMLF-IFD under bidirectional power flow conditions, qualitative analysis and quantitative evaluation are presented. As for qualitative analysis, we considered the bidirectional power flow in the distribution system and included corresponding instances in the training dataset. Through the training process, FSMLF-IFD can learn the characteristics of this bidirectional power flow and achieve IF detection and identification under bidirectional power flow conditions. As for quantitative evaluation, we present two cases to verify that FSMLF-IFD is effective under bidirectional power flow conditions.

- Case 1 (forward power flow): The load is set to 15.06 MW, and the wind farm's output is 4 MW. At this time, the power flow is from the grid to the load, which is the forward power flow.
- Case 2 (reverse power flow): The load is set to 11.06 MW, and the wind farm's output is 6 MW. At this time, the power flow is from the load to the grid, which is the reverse power flow.

Table 4 shows the IF identification accuracy of the proposed method under the conditions of the two cases. In both cases, the proposed method achieved good performance and consistency. This indicates that the proposed method is effective under bidirectional power flow conditions.

In addition, for multi-cycle IFs, we consider dividing the multi-cycle IF into multiple sub-cycle faults. If a fault occurs in multiple consecutive cycles, this fault is determined as a multi-cycle fault. Based on this idea, even if the fault is not detected in the first cycle, it can be detected by subsequent cycles. To further demonstrate the effectiveness of this idea for FSMLF-IFD, the confusion matrix for detecting multi-cycle IFs is shown in Fig. 9. For 10 multi-cycle IFs, only one fault event is not detected, which proves the effectiveness of the proposed detection strategy.

D. EFFECTIVE ADAPTIVE IMAGE FUSION METHOD

The adaptive image fusion method is essential for our proposed framework. The weight of merging the current waveform image and the voltage waveform image is calculated using the adaptive image fusion strategy. Fig. 10 illustrates the performance comparison of various fusion

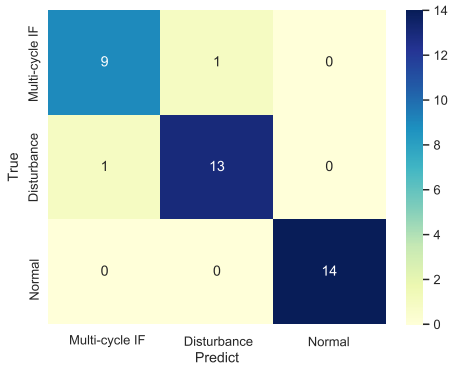


FIGURE 9. Confusion matrix for detection multi-cycle IFs. The value in this figure represents the number of current events.

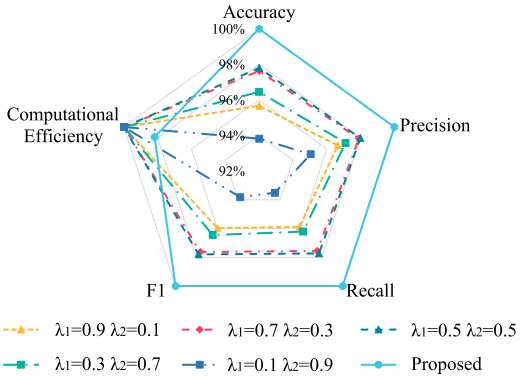


FIGURE 10. Performance comparison diagram of different fusion weights.

weights for IF detection, where λ_1 and λ_2 denote the fusion weights of the current image and voltage image, respectively. Although the adaptive fusion method increases the computational burden, the performance of the fusion method used in this paper is significantly superior to the fixed fusion weight under the four evaluation indicators. When λ_1 and λ_2 are 0.5 respectively, the fixed fusion weight achieves the best performance. However, it is still 2.1% worse than the classification accuracy of the adaptive fusion strategy, demonstrating the superiority of this method in performance. Additionally, when $\lambda_1 = 0.1$ and $\lambda_2 = 0.9$, the classification accuracy after fusion is only 93.8% of the adaptive fusion strategy. It can be seen from Fig. 10 that the fusion parameters are sensitive to the detection method based on waveform images. Therefore, the adaptive fusion strategy is necessary and effective.

E. ADVANTAGES BASED ON ADAPTABILITY-ENHANCING WEIGHTING INITIALIZATION STRATEGY

We will demonstrate the advantages of the adaptability-enhancing weighting initialization strategy of FSMLF-IFD based on real IF data. Table 5 presents the IF detection results of different shots with and without the initialization strategy. It is worth noting that the results presented in the table are the average of multiple experiments. The results shown in the table are based on the performance of adaptive image fusion. Firstly, the accuracy of FSMLF-IFD can

TABLE 5. Performance comparison of different shots with and without the initialization strategy for IF detection.

Shot	Method	Accuracy	Precision	Recall	F1
1-shot	without	0.8994	0.9214	0.9002	0.9106
	with	0.9487	0.9585	0.9497	0.9541
3-shot	without	0.9414	0.9536	0.9395	0.9465
	with	0.9720	0.9760	0.9843	0.9801
5-shot	without	0.9570	0.9634	0.9582	0.9607
	with	0.9746	0.9775	0.9850	0.9811

Note: "with" indicates that the adaptability-enhancing weighting initialization strategy is applied to the modeling process. "without" indicates that the adaptability-enhancing weighting initialization strategy is not applied to the modeling process.

achieve 94.87% with 1-shot, which shows the effectiveness of the proposed framework. Secondly, compared with the meta-learning without initialization strategy, the performance of FSMLF-IFD has improved by 5.19%, 3.24%, 1.80% under 1-shot, 3-shot, and 5-shot learning, respectively. As a result, the adaptability-enhancing weighting initialization strategy ensure that FSMLF-IFD is specifically designed for IF detection.

F. COMPLEXITY AND SENSITIVITY ANALYSES OF FSMLF-IFD

In this subsection, the complexity and sensitivity analyses of FSMLF-IFD are provided. The computational complexity is important for detection method. To evaluate the computational requirements of the IF detection method, the big O notation [40] and detection time are considered. Because input image sizes and sampling rate of signals have a crucial impact on estimation accuracy, the sensitivity of FSMLF-IFD for two parameters is analyzed.

The big O notation of the IF detection process for FSMLF-IFD and two state-of-the-art methods is discussed. In Hybrid-SVM [6], ST and SVM are used for IF detection. The time complexity of the ST is $O(N^2)$, where N is the length of the input data and $O(\cdot)$ represents big O notation. Considering the Gaussian kernel, the complexity of SVM is $O\left(N^2\right)$. In sum, the overall complexity of [6] is $O(N^2) + O(N^2) \approx \mathrm{O}(N^2)$. For FSMLF-IFD and TSML [32], CNNs are used for IF detection. In the forward propagation process, the computational load mainly comes from convolution and pooling operations. The time complexity of the CNNs is approximate $O(LKMCS^2)$, where L is the number of layers, K is the number of convolution kernels, C is the size of the channel, M is the size of convolution kernels, S is the size of the waveform image. According to the parameters set in this article, the obtained result of complexity is $O(108S^2)$. For the waveform generation strategy, the complexity is approximately equal to $O(4S^2)$. Then, the complexity of FSMLF-IFD is still $O(108S^2) + O(4S^2) \approx \mathrm{O}(112S^2)$. Due to $O(112S^2) > O(N^2)$, FSMLF-IFD and TSML have higher computational requirements compared to [6].

TABLE 6. Comparison of computational requirement.

Detection model	Big O notation	Detection time
Hybrid-SVM [6]	$O(N^2)$	0.0008 s
TSML [32]	$O(108S^2)$	0.0010 s
FSMLF-IFD	$O(112S^2)$	0.0012 s

Next, we tested the detection time of three methods on the current experimental equipment. For a detection process, all three methods are very fast. The time required for detection by FSMLF-IFD, TSML, and Hybrid-SVM is approximately 0.0012 s, 0.0010 s, and 0.0008 s, respectively. Although FSMLF-IFD takes the most time, the differences compared to the other two methods are negligible.

The comparison of computational requirements is presented in Table 6. The discussion from the two perspectives presented above leads to the conclusion that FSMLF-IFD requires slightly more computational demands compared to two state-of-the-art methods, however, this discrepancy is inconsequential.

Table 7 shows the performance comparison results of different image sizes and signal sampling rates. For different image sizes, the performance exhibited by FSMLF-IFD varies to some extent. When the image size is 32×32, the accuracy of FSMLF-IFD can still achieve 0.7916. When the sizes of images are higher than 64×64, the accuracy of FSMLF-IFD can be greater than 0.8625, which proves that the proposed method is insensitive to image size. As seen in Table 7, at different sampling rates, FSMLF-IFD exhibits varying performance. As the sampling rate increases, the performance gradually improves. At a sampling rate of 1.6 kHz, FSMLF-IFD can still achieve an accuracy of 0.8222. This indicates the performance of FSMLF-IFD is not sensitive to changes in the sampling rate. In addition, it is important to emphasize that the sampling rate for FSMLF-IFD is not limited to 6.4kHz. The sampling rate of 6.4kHz is not a necessity for the proposed method and 6.4kHz is merely an example used in our case study. Of course, the proposed method is applicable to data with various sampling rates, depending on the sampling rate of the waveform data collection equipment.

G. COMPARISON WITH COMMERCIAL PRODUCTS FOR IF DETECTION

To furthermore verify the advantages of the proposed method, comprehensive comparison results with two commercial products are presented in Table 8. The two commercial products are Early Fault Detection (EFD) [41], [42] and Distribution Fault Anticipation (DFA) [43].

The red area represents the drawbacks of the two products. For the product EFD, the data is based on current and voltage

TABLE 7. The performance comparison results of different image sizes and signal sampling rate on the realistic dataset.

Condition	Parameters	Accuracy	Precision	Recall	F1
Sampling rate	12.8kHz	0.9746	0.9775	0.9840	0.9816
	6.4kHz	0.9720	0.9760	0.9843	0.9801
	3.2kHz	0.9354	0.9467	0.9388	0.9427
	1.6kHz	0.8222	0.8364	0.8276	0.8320
Image size	256*256	0.9720	0.9760	0.9843	0.9801
	128*128	0.9488	0.9550	0.9520	0.9534
	64*64	0.8625	0.8993	0.8667	0.8826
	32*32	0.7916	0.8025	0.8416	0.8215

TABLE 8. Comparison results between FSMLF-IFD and two commercial products.

Method or product	Data	Data volume	Algorithm	Detection time
EFD	Current and voltage traveling waves	Unknow (patented)	Traveling wave method and pattern recognition	Real-time
DFA [43]	Current and voltage waveform	Large sample size	Fuzzy expert system, fuzzy dynamic time warping, etc	Within 3 min
FSMLF-IFD	Current and voltage waveform	Small sample size	Waveform image generation and meta-learning framework	0.001s

traveling waves. The algorithm is based on the traveling wave algorithm and pattern recognition. The basic principle of this product is that failing equipment emits the traveling wave and some specific measuring devices are used to capture these traveling waves. Then, the obtained current and voltage traveling wave data was analyzed using the traveling wave algorithm. According to the results of online searches, the traveling wave algorithm is based on the frequency-time signature analysis method. Finally, the traveling waves are based on pattern recognition to translate into condition monitoring information such as 'equipment type' and 'location' [41]. However, in the signal transmission process, radio frequency sensors are needed for communication among the various receivers of the power line signal [42]. This indicates the product requires additional equipment and the cost of such equipment is high. In contrast, our proposed method focuses on analyzing current and voltage waveforms, which can be obtained from substation-based feeder CTs and bus PTs, etc., thus avoiding high costs.

As for product DFA, its drawback is that large amounts of data and very complex computational algorithms. These drawbacks resulted in longer detection times. The algorithm of DFA is based on traditional learning methods, such as expert systems. However, these algorithms lack intelligence and are outdated. Compared with DFA, this paper proposes a waveform image generation strategy and a few-shot meta-learning framework. The greatest advantage is the ability to learn from small samples.

It is worth noting that the above analysis does not indicate that the method referred to is stronger than the existing products, we only prove the proposed method is outstanding in some respects.

VI. CONCLUSION

This paper focuses on learning power systems waveform IF patterns through an intelligent strategy. A few-shot learning framework, FSMLF-IFD, is presented to achieve learning incipient patterns with only a few samples. This framework provides prior knowledge for the IF learner by mining current and voltage waveform images of other fault types and is helpful to power system protection and the development of smart grids. To obtain input images of FSMLF-IFD, the waveform image conversion strategy and adaptive image fusion strategy are proposed. These two strategies integrate voltage and current waveform data information into comprehensible images. The comprehensible image can capture the time-varying characteristics of amplitude and waveform shapes. To address data differences during training the FSMLF-IFD, an adaptability-enhancing weighting initialization strategy is developed. This strategy enhances adaptability and accuracy of IF detection model. In the numerical results, the superiority and efficiency of FSMLF-IFD are verified by the field IF data and simulated distribution systems with renewable energy sources. However, the IF detection models based on data-driven methods are susceptible to factors such as changes in power flow and topology. In future work, the impact of bidirectional flow and the solutions to address changes in system parameters will be studied.

REFERENCES

[1] L. M. Bieber, L. Wang, and W. Li, "A low-loss thyristor-based hybrid three-level and modular multilevel converter with DC fault blocking capability for HVDC transmission," *IEEE Open Access J. Power Energy*, vol. 7, pp. 111–121, 2020.

[2] M. Parsi and P. A. Crossley, "Optimised time for travelling wave fault locators in the presence of different disturbances based on real-world fault data," *IEEE Open Access J. Power Energy*, vol. 8, pp. 138–146, 2021.

[3] G. N. Lopes, T. S. Menezes, D. P. S. Gomes, and J. C. M. Vieira, "High impedance fault location methods: Review and harmonic selection-based analysis," *IEEE Open Access J. Power Energy*, vol. 10, pp. 438–449, 2023.

[4] B. Chen, "Fault statistics and analysis of 220-kV and above transmission lines in a southern coastal provincial power grid of China," *IEEE Open Access J. Power Energy*, vol. 7, pp. 122–129, 2020.

[5] M. Bin Gani and S. Brahma, "A closed-form mathematical model and method for fast fault location on a low voltage DC feeder using single-ended measurements," *IEEE Open Access J. Power Energy*, vol. 9, pp. 523–536, 2022.

[6] G. W. Chang, Y.-H. Hong, and G.-Y. Li, "A hybrid intelligent approach for classification of incipient faults in transmission network," *IEEE Trans. Power Del.*, vol. 34, no. 4, pp. 1785–1794, Aug. 2019.

[7] X. Dong, Y. Huang, H. Wang, B. Chen, H. Wang, and Q. Dong, "Analysis and simulation research of cascading faults in AC/DC hybrid grid," *IEEE Open Access J. Power Energy*, vol. 9, pp. 514–522, 2022.

[8] P. Q. Subcommittee, "Electric signatures of power equipment failures," IEEE, USA, Tech. Rep., PES-TR73, Dec. 2019.

[9] H. Mohsenian-Rad, *Smart Grid Sensors: Principles and Applications*. Cambridge, U.K.: Cambridge Univ. Press, 2022.

[10] M. Jannati, B. Vahidi, and S. H. Hosseinian, "Incipient faults monitoring in underground medium voltage cables of distribution systems based on a two-step strategy," *IEEE Trans. Power Del.*, vol. 34, no. 4, pp. 1647–1655, Aug. 2019.

[11] J. A. Wischkaemper, C. L. Benner, B. D. Russell, and K. Manivannan, "Application of waveform analytics for improved situational awareness of electric distribution feeders," *IEEE Trans. Smart Grid*, vol. 6, no. 4, pp. 2041–2049, Jul. 2015.

[12] H. Mohsenian-Rad and W. Xu, "Synchro-waveforms: A window to the future of power systems data analytics," *IEEE Power Energy Mag.*, vol. 21, no. 5, pp. 68–77, Sep. 2023.

[13] X. Jiang, B. Stephen, and S. McArthur, "A sequential Bayesian approach to online power quality anomaly segmentation," *IEEE Trans. Ind. Informat.*, vol. 17, no. 4, pp. 2675–2685, Apr. 2021.

[14] Y. Weng, Q. Cui, and M. Guo, "Transform waveforms into signature vectors for general-purpose incipient fault detection," *IEEE Trans. Power Del.*, vol. 37, no. 6, pp. 4559–4569, Dec. 2022.

[15] Q. Li et al., "Incipient fault detection in power distribution system: A time–frequency embedded deep-learning-based approach," *IEEE Trans. Instrum. Meas.*, vol. 72, pp. 1–14, 2023.

[16] S. Kulkarni, S. Santoso, and T. A. Short, "Incipient fault location algorithm for underground cables," *IEEE Trans. Smart Grid*, vol. 5, no. 3, pp. 1165–1174, May 2014.

[17] T. S. Sidhu and Z. Xu, "Detection of incipient faults in distribution underground cables," *IEEE Trans. Power Del.*, vol. 25, no. 3, pp. 1363–1371, Jul. 2010.

[18] M. F. Faisal, A. Mohamed, and H. Shareef, "Prediction of incipient faults in underground power cables utilizing S-transform and support vector regression," *Int. J. Electr. Eng. Informat.*, vol. 4, no. 2, pp. 186–201, Jun. 2012.

[19] S. Xiong, Y. Liu, J. Fang, J. Dai, L. Luo, and X. Jiang, "Incipient fault identification in power distribution systems via human-level concept learning," *IEEE Trans. Smart Grid*, vol. 11, no. 6, pp. 5239–5248, Nov. 2020.

[20] M. Izadi and H. Mohsenian-Rad, "A synchronized Lissajous-based method to detect and classify events in synchro-waveform measurements in power distribution networks," *IEEE Trans. Smart Grid*, vol. 13, no. 3, pp. 2170–2184, May 2022.

[21] Y. Wu, P. Zhang, and G. Lu, "Detection and location of aged cable segment in underground power distribution system using deep learning approach," *IEEE Trans. Ind. Informat.*, vol. 17, no. 11, pp. 7379–7389, Nov. 2021.

[22] S. Wang and H. Chen, "A novel deep learning method for the classification of power quality disturbances using deep convolutional neural network," *Appl. Energy*, vol. 235, pp. 1126–1140, Feb. 2019.

[23] A. Bagheri, I. Y. H. Gu, M. H. J. Bollen, and E. Balouji, "A robust transform-domain deep convolutional network for voltage dip classification," *IEEE Trans. Power Del.*, vol. 33, no. 6, pp. 2794–2802, Dec. 2018.

[24] S. Wang and P. Dehghanian, "On the use of artificial intelligence for high impedance fault detection and electrical safety," *IEEE Trans. Ind. Appl.*, vol. 56, no. 6, pp. 7208–7216, Nov. 2020.

[25] S. Ekici, F. Ucar, B. Dandil, and R. Arghandeh, "Power quality event classification using optimized Bayesian convolutional neural networks," *Electr. Eng.*, vol. 103, no. 1, pp. 67–77, Feb. 2021.

[26] M. Izadi and H. Mohsenian-Rad, "Characterizing synchronized Lissajous curves to Scrutinize power distribution synchro-waveform measurements," *IEEE Trans. Power Syst.*, vol. 36, no. 5, pp. 4880–4883, Sep. 2021.

[27] R. S. Salles and P. F. Ribeiro, "The use of deep learning and 2-D wavelet scalograms for power quality disturbances classification," *Electric Power Syst. Res.*, vol. 214, Jan. 2023, Art. no. 108834.

[28] J. Shukla, B. K. Panigrahi, and P. K. Ray, "Power quality disturbances classification based on Gramian angular summation field method and convolutional neural networks," *Int. Trans. Electr. Energy Syst.*, vol. 31, no. 12, Dec. 2021, Art. no. e13222.

[29] P. Hart et al., "Application of big data analytics and machine learning to large-scale synchrophasor datasets: Evaluation of dataset 'machine learning-readiness,'" *IEEE Open Access J. Power Energy*, vol. 9, pp. 386–397, 2022.

[30] A. G. C. Menezes, M. M. Araujo, O. M. Almeida, F. R. Barbosa, and A. P. S. Braga, "Induction of decision trees to diagnose incipient faults in power transformers," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 29, no. 1, pp. 279–286, Feb. 2022.

[31] C. A. Andresen, B. N. Torsæter, H. Haugdal, and K. Uhlen, "Fault detection and prediction in smart grids," in *Proc. IEEE 9th Int. Workshop Appl. Meas. Power Syst. (AMPS)*, Sep. 2018, pp. 1–6.

[32] Y. Hu, R. Liu, X. Li, D. Chen, and Q. Hu, "Task-sequencing meta learning for intelligent few-shot fault diagnosis with limited data," *IEEE Trans. Ind. Informat.*, vol. 18, no. 6, pp. 3894–3904, Jun. 2022.

[33] S. Zhang, F. Ye, B. Wang, and T. G. Habetler, "Few-shot bearing fault diagnosis based on model-agnostic meta-learning," *IEEE Trans. Ind. Appl.*, vol. 57, no. 5, pp. 4754–4764, Sep. 2021.

[34] S. Kulkarni, A. J. Allen, S. Chopra, S. Santoso, and T. A. Short, "Waveform characteristics of underground cable failures," in *Proc. IEEE PES Gen. Meeting*, Jul. 2010, pp. 1–8.

[35] W. Zhang, Y. Jing, and X. Xiao, "Model-based general arcing fault detection in medium-voltage distribution lines," *IEEE Trans. Power Del.*, vol. 31, no. 5, pp. 2231–2241, Oct. 2016.

[36] C. Finn, P. Abbeel, and S. Levine, "Model-agnostic meta-learning for fast adaptation of deep networks," in *Proc. 34th Int. Conf. Mach. Learn.*, Aug. 2017, pp. 1126–1135.

[37] Q. Cui, K. El-Arroudi, and Y. Weng, "A feature selection method for high impedance fault detection," *IEEE Trans. Power Del.*, vol. 34, no. 3, pp. 1203–1215, Jun. 2019.

[38] M. Kizilcay and K. H. Koch, "Numerical fault arc simulation based on power arc tests," *Eur. Trans. Electr. Power*, vol. 4, no. 3, pp. 177–185, May 1994.

[39] M. Oleskovicz, D. V. Coury, O. D. Felho, W. F. Usida, A. A. F. M. Carneiro, and L. R. S. Pires, "Power quality analysis applying a hybrid methodology with wavelet transforms and neural networks," *Int. J. Electr. Power Energy Syst.*, vol. 31, no. 5, pp. 206–212, Jun. 2009.

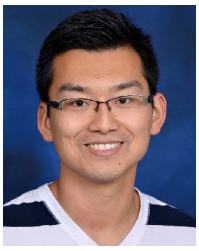
[40] S. G. Devi, K. Selvam, and S. Rajagopalan, "An abstract to calculate big O factors of time and space complexity of machine code," in *Proc. Int. Conf. Sustain. Energy Intell. Syst.*, 2011, p. 844.

[41] *IND-EFD Installed on VIC Peninsula*. Accessed: 2021. [Online]. Available: https://ind-technology.com.au/efd-installed-on-vic-peninsula

[42] *Incubatenergy Labs 2020 Pilot Project Report: IND Technology—Early Fault Detection for Power Lines*. Accessed: 2021. [Online]. Available: https://www.epri.com/research/products/000000003002020659

[43] B. Don Russell, C. Benner, J. Wischkaemper, and K. Muthu-Manivannan, "Incipient electric circuit failure detection and outage prevention using advanced electrical waveform monitoring: Field experience," *IEEE Ind. Appl. Mag.*, vol. 29, no. 3, pp. 36–45, May 2023.

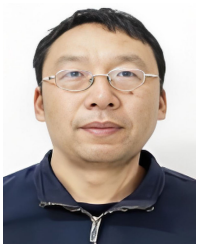
LIXIAN SHI received the M.S. degree in instrument engineering from Kunming University of Science and Technology, Kunming, China, in 2021. He is currently pursuing the Ph.D. degree in electrical engineering with Chongqing University. His work centers on incipient fault detection and power quality disturbance detection and identification.



QIUSHI CUI (Member, IEEE) received the M.Sc. degree in electrical engineering from Illinois Institute of Technology and the Ph.D. degree in electrical engineering from McGill University. He was a Post-Doctoral Researcher with Arizona State University (ASU) and the Associate Director of the Machine Learning Laboratory for Power Systems, Ira A. Fulton Schools of Engineering, Chongqing University. Prior to joining ASU, he was a Research Engineer with OPAL-RT Technologies Inc., from November 2015 to November 2017. Currently, he is an Associate Professor with the School of Electrical Engineering, Chongqing University. His research interests include machine learning and big data applications in power systems, power system protection, smart cities, microgrid, EV integration, and real-time simulation in power engineering. He has won three best paper awards from U.K., China, and the USA, all ranking the first. He was the Winner of the Chunhui Cup Innovation and Entrepreneurship Competition for Overseas Chinese Scholars in the Energy Sector in 2018. He received the Post-Doctoral Research Scholarship from both Natural Sciences and Engineering Research Council of Canada (NSERC) and Quebec Research Fund-Nature and Technology (FRQNT) and held the MITACS Accelerate Research Program Fellowship from Canada.

YANG WENG (Senior Member, IEEE) received the B.E. degree in electrical engineering from Huazhong University of Science and Technology, Wuhan, China, the M.Sc. degree in statistics from the University of Illinois at Chicago, Chicago, IL, USA, the M.Sc. degree in machine learning of computer science, and the M.E. and Ph.D. degrees in electrical and computer engineering from Carnegie Mellon University (CMU), Pittsburgh, PA, USA. He joined Stanford University, Stanford, CA, USA, as a TomKat Fellow of Sustainable Energy. He is currently an Assistant Professor in electrical, computer, and energy engineering with Arizona State University, Tempe, AZ, USA. His research interests include power systems, machine learning, and renewable integration. He was a recipient of the CMU Deans Graduate Fellowship in 2010, the Best Paper Award at the International Conference on Smart Grid Communication (SGC) in 2012, the First Ranking Paper of SGC in 2013, the Best Papers at the Power and Energy Society General Meeting in 2014, the ABB Fellowship in 2014, and the Golden Best Paper Award at the International Conference on Probabilistic Methods Applied to Power Systems in 2016.

YIGONG ZHANG (Student Member, IEEE) received the B.S. and M.S. degrees from Chongqing University, Chongqing, China, in 2018 and 2021, respectively. He is currently pursuing the Ph.D. degree in electrical engineering with a focus on multienergy microgrids and artificial intelligence methods in power systems.



SHILONG CHEN received the Ph.D. degree from Kunming University of Science and Technology, Kunming, China, in 2012. He is currently a Professor and a Master's Supervisor with the Faculty of Electric Power Engineering, Kunming University of Science and Technology. He has presided over three projects funded by the National Natural Science Foundation of China. His main research interests include HVDC transmission and new relay protection for power systems.

JIAN LI (Senior Member, IEEE) received the M.S. and Ph.D. degrees in electrical engineering from Chongqing University, Chongqing, China, in 1997 and 2001, respectively. He is currently the Vice President of Chongqing University. His major research interests include the intelligence energy, the Internet of Things, online detection of insulation condition in electrical devices, partial discharge, and insulation fault diagnosis for high-voltage equipment.

WENYUAN LI (Life Fellow, IEEE) is currently a Professor with Chongqing University, Chongqing, China. His research interests include power system planning, operation, optimization, and reliability assessment. He is a fellow of Canadian Academy of Engineering and the Engineering Institute of Canada and a Foreign Member of Chinese Academy of Engineering. He was a recipient of several IEEE PES awards, including the IEEE PES Roy Billinton Power System Reliability Award in 2011 and the IEEE Canada Electric Power Medal in 2014.