IET Generation, Transmission & Distribution

WILEY

The Institution of Engineering and Technology

ORIGINAL RESEARCH OPEN ACCESS

An AI-Based Technique for Fault Location in Inverter-Based Active Distribution Networks

Morteza Behbahanipour[1] | Seyed Fariborz Zarei[1] | Mohammadhadi Shateri[2]

[1]Department of Electrical and Computer Engineering, Qom University of Technology, Qom, Iran | [2]Department of Systems Engineering, École De Technologie Supérieure, Montreal, Quebec, Canada

Correspondence: Seyed Fariborz Zarei (zarei@qut.ac.ir)

Received: 19 September 2025 | **Revised:** 4 December 2025 | **Accepted:** 20 December 2025

ABSTRACT

Fault location is increasingly essential in inverter-based active distribution networks. This is due to the large number of branches and laterals in such networks, as well as the presence of inverter-based distributed generators (IBDGs). Several techniques are used for locating faults in distribution networks, including impedance-based approaches, traveling wave-based schemes, and artificial intelligence (AI)-based approaches. AI-based schemes are superior to others in terms of speed and accuracy, and they do not demand high-frequency devices. However, there is a lack of AI-based schemes that can effectively address scenarios involving a high number of branches, a limited number of measurement instruments, the presence of IBDGs, and high fault resistance. Accordingly, this paper introduces a modified one-dimensional convolutional neural network (1-D CNN) that combines residual connections with 1-D CNNs. The suggested approach includes two elements for fault location: (i) determining the fault distance and (ii) identifying the section of the network that is faulty. The results indicate that this approach effectively pinpoints faults with varying resistance levels at different locations, even in the presence of IBDGs. Ultimately, the proposed solution demonstrates enhanced accuracy in networks featuring multiple distributed generators, numerous sub-branches, unbalanced load conditions, and diverse fault scenarios.

1 | Introduction

Accurately identifying faults in power distribution systems is essential for their effective operation. Quick fault detection helps restore service to affected areas and enables the faulty line to resume functioning [1]. Fault location methods are designed to provide a more reliable and efficient approach to detecting and addressing faults in distribution networks [2]. This algorithm is responsible for identifying both the section of the network that has faulted and the specific location of the fault in that section. The complexity of fault location is heightened by the numerous interconnected feeders and laterals in the distribution network. Additionally, the fault resistance can significantly influence the detection process. The integration of an inverter-based distributed generator (IBDG) into the network adds another layer of complexity [3]. This integration transforms the network into a multi-source system with various feeding points, complicating the fault location process and requiring more thorough analysis. During short-circuit conditions, managing the output currents from inverters becomes critical, further complicating the behaviour of IBDGs and impacting fault detection [4]. These factors can hinder the effectiveness of fault detection methods.

Fault location techniques are generally categorized into three types of: (i) impedance-based approaches, (ii) traveling wave-based schemes and (iii) artificial intelligence-based approaches

Abbreviations: 1-D CNN, one-dimensional convolutional neural network; ANN, artificial neural network; CNN, convolutional neural network; DGs, distributed generators; IBDG, inverter-based distributed generation; KNN, *k*-nearest neighbours; LSTM, long short-term memory; SVM, support vector machine.

[5]. Artificial intelligence-based methods have much higher speed and accuracy than other methods in the subject of fault location. The fault location accuracy in these methods is often above 95%. Unlike impedance-based methods, these methods do not have problems such as multiple answers and dependence on network information and loads and lines [6]. Also, unlike traveling wave-based schemes, these methods do not require high-frequency devices. Due to these reasons, various articles and research studies have focused on artificial intelligence-based methods in recent years [7].

Methods that utilize artificial intelligence commonly employ various tools, including CNNs, SVMs and graph neural networks, for the purpose of identifying faults. These techniques will be discussed in detail in the following paragraphs.

Several studies leverage convolutional neural networks for fault classification and localization. A method [8] transforms current signals to images and processes them with a CNN framework to identify the fault type and section. Simulations on an 8-bus network achieved results with a 100 Ω fault resistance. A one-dimensional convolutional neural network (1-D CNN) with waveform concatenation [9] adaptively extracts features from the transient of the zero-sequence currents to locate the fault section. The method achieved 93% accuracy on an 11-bus, 10 kV radial network. Another study [10] introduces a unique method for detecting line-to-ground faults in distribution networks. This technique combines a patch-to-patch CNN with a feeder-to-feeder LSTM network. The CNN effectively extracts deep spatial features, which are then processed by the LSTM to maintain the temporal relationships of these features. The proposed patch-to-patch CNN recognizes current waveforms at a granular level, while the feeder-to-feeder LSTM learns and compares spatial-temporal correlations across feeders. The integration of these two networks enhances the ability to identify fault characteristics in zero-sequence currents. Testing on a 10 kV radial network with nine buses, six measuring devices, and a fault resistance of 2000 Ω yielded an impressive accuracy of 98.8%. A fast hybrid method [11] combines an ANN to estimate the distances from measurement points, and a KNN classifier to interpret the ANN outputs and pinpoint the fault location. This technique uses voltage-current data from the main substation, DG connection points, and microgrids.

Graph neural networks are emerging as powerful tools for fault detection. A temporal recurrent graph neural network in [12] extracts spatial-temporal features from voltage measurement data to perform fault event detection, classification and location. Simulations on a 13.2 kV, 123-bus ring network with 10 Ω fault resistance achieved 91.6% accuracy. A heterogeneous multi-task learning graph neural network can detect, locate and classify faults, while providing fault resistance and current estimates. Simulations on a 123-bus, 4.1 kV radial network with 500 Ω fault resistance yielded 93% accuracy [13]. Deep graph learning models consider the physical network structure as a constraint to enhance abnormal data resistance and generalization. Simulations on a 33-bus, 12.6 kV radial network with 11 measuring devices and 600 Ω fault resistance yielded 99% accuracy [14].

The SVM-based techniques are also considered as machine learning methods. Multi-task learning with SVM classification and SVR regression is used in [15] to detect the class of the fault, faulted line and predict the precise fault location. Simulations on a 9-bus radial network with nine measuring devices and four DGs achieved 98.99% accuracy. The [16] utilizes an SVM for recognizing the condition of power equipment. This process involves collecting electrical data, preprocessing it, extracting relevant features, and classifying the information. Additionally, fault-traveling wave signals are analysed through variation mode decomposition to determine the fault interval and calculate the precise distance to the fault. This approach was validated on a 138 kV, 24-bus network featuring 11 DGs. Gaussian process regression and SVMs are used for fault location prediction and identification, respectively, based on measuring one-cycle RMS voltages and currents. The error was 0.28 on a 7-bus, 33 kV network with three DGs [17].

Other machine learning methods have also been explored. Deep reinforcement learning can identify fault sections without relying on large labelled datasets. The methodology [18] was tested on the IEEE 14-bus system with five DGs. Variable mode decomposition, t-SNE manifold learning and fuzzy clustering are used to de-noise transient zero-sequence currents, visualize fault information and determine the fault section. Simulations on a 13-bus, 10 kV radial network with 3000 Ω fault resistance and two DGs were conducted [19]. Approach [20] uses gradient boosting trees to localize line-to-ground and balanced faults in a low-voltage system. The method assumes general branch-independent features across grid topologies. Simulations on a 32-bus radial network with 18 DGs achieved an 84% accuracy in determining the fault section.

The techniques discussed earlier face several challenges, such as limited fault resistance, a distribution network with few sub-branches, insufficient integration of digital meters, and the extensive distribution of numerous meters throughout the network. To tackle these issues effectively, we employed a one-dimensional conventional neural network model that incorporates a residual block. This approach aims to accurately pinpoint fault locations and determine the distance to these faults. Here are some key contributions and novelties regarding our methodology.

- A larger and more branched distribution network is selected for testing, with practical meter placement and appropriately incorporated distributed generation to reflect real operating conditions.
- Short-circuit faults with fault-resistance are examined to evaluate performance under more challenging scenarios.
- Both faulted-section identification and fault-distance estimation are addressed simultaneously, as both are essential for accurate fault location.
- A residual 1-D CNN architecture is employed to extract weak and nonlinear transient features from faults with Rf and IBDG dynamics.
- Residual (skip) connections are used to maintain gradient flow across deeper layers, preserving subtle fault signatures.
- A multi-task learning framework shares a common feature extractor for both tasks, improving generalization under limited observability.

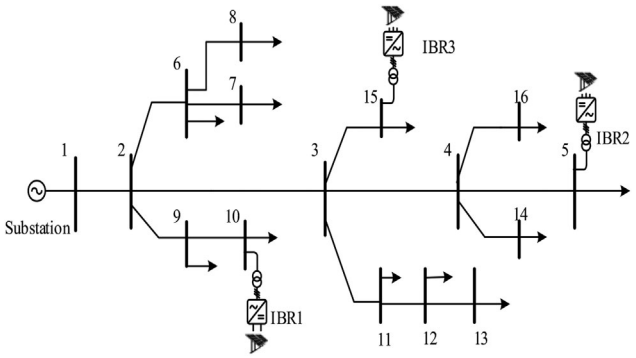

FIGURE 1 | Schematic of the used test system.

TABLE 1 | IBDGS data of test system.

Parameters	Values
Transformers	Nominal voltage: 0.25 KV/11 KV
	Nominal power: 250 KVA
	R1 (pu), L1 (pu): 0.0012, 0.03
PVs	Sun Power SPR-415E
	Parallel strings: 88
	Series-connected modules per string: 7
	Irradiance: 1000 W/m^2
	Temperature: 45°C
Inverters	Nominal power: 250 kVA
	Primary voltage: 11 kV
	Secondary voltage: 0.25 kV
	DC voltage: 480 V

The rest of the paper is organized as follows: In Section 2, we will outline the testing system and the factors used for data collection. Section 3 focuses on the fundamental principles of the suggested fault location technique, which is based on a traditional one-dimensional neural network. This section will also describe the proposed method for pinpointing the faulty segment and measuring the distance to the fault. In Section 4, we will present the simulation results of our model, including a detailed comparison with other artificial intelligence-based approaches. Finally, Section 5 will conclude with a summary of the main findings and contributions of this research.

2 | Test System and Data Gathering

The single-line diagram of a modified IEEE 15-bus radial distribution system is presented in Figure 1. This three-phase, 11 kV, 50 Hz system is based on line parameters and load data sourced from [21]. Three photovoltaic inverters, acting as IBDGs, are integrated at buses 5, 10, and 15. With a total of 16 buses in the chosen test system, incorporating three IBDGs poses difficulties for locating faults. This observation aligns with the strict conditions outlined at the end of Section 1. Detailed information about the IBDGs is listed in Table 1. To evaluate the proposed scheme, detailed time-domain simulations were conducted in the MATLAB/Simulink environment using this modified 15-bus distribution feeder.

The system consists of 16 buses, 14 line sections, 12 load buses, and three IBDGs with a total three-phase power capacity of 7.5 MVA. Each IBDG contributes 250 kVA and operates at an output voltage of 250 V, connected to the network via a transformer [22]. A four-wire grounded neutral RL model was employed to validate the methodology, as the feeder configuration exhibits unequal distances between phases and non-transposed lines, resulting in an unbalanced line impedance matrix [23]. The impedance of the distribution network lines and the power of the loads connected to their ends are listed in Table 2. The line impedance matrix accounts for the unbalanced nature of the system due to the unequal distances between phases and non-transposed lines. The load data for each bus is also provided, allowing for a comprehensive understanding of the system's power consumption and distribution.

It is worth mentioning that the four-wire grounded-neutral RL modelling used in this paper ensures that the simulated three-phase voltage and current waveforms retain the unequal mutual coupling and phase-to-phase asymmetries inherent to non-transposed lines. Consequently, the measured signals are represented as multi-channel time series ($X \in \mathbb{R}^{T\times D}$), where T is the number of time samples and D the number of measured channels. These signals inherently encode the inter-phase coupling patterns produced by the unbalanced impedance matrix. Physically, the unequal mutual inductances introduce cross-phase transient components and sequence mixing during faults; these effects manifest as phase-dependent waveform distortions and distinctive temporal correlations between channels, which differ significantly from those observed in ideally transposed systems. Because conventional impedance- or sequence-based methods often rely on symmetrical-component transformations that assume line transposition—and therefore may yield biased or incomplete representations in non-transposed feeders—the raw multi-channel waveforms in this dataset are retained so that the data-driven algorithms (Section 3) can learn the pertinent inter-phase coupling signatures directly. The simulated scenarios include both balanced and the full range of unbalanced short-circuit types, ensuring that the dataset captures the complete spectrum of symmetrical and asymmetrical transient behaviours required to evaluate performance under realistic, unbalanced operating conditions.

2.1 | Data Gathering

In our proposed approach, the neural network utilizes three-phase voltage and current data collected from the start of the distribution network (substation) and from buses linked to DGs. Based on the analysed network depicted in Figure 1, we gathered three-phase voltage and current readings from buses 1, 5, 10, and 15. The selection of these buses for measurement is motivated by the practical constraint that measurements are not available at every point in the network, and the installation of additional measurement devices across many locations is neither practical nor cost-effective. In typical distribution systems, measurements are available at the substation (busbar 1 in Figure 1) and at the connection points of distributed generators, and these existing

TABLE 2 | Line and load data of test system.

Line No.	From	To	Line data		Load data (at destination bus)	
			r (Ω)	x (Ω)	kW	kVar
1	1	2	1.35309	1.32349	44.1	44.990
2	2	3	1.17024	1.14464	70.1	71.440
3	3	4	0.84111	0.82271	40.0	142.820
4	4	5	1.52348	1.0276	44.1	44.990
5	2	9	2.01317	1.3279	70.0	71.440
6	9	10	1.68671	1.1377	44.1	44.990
7	2	6	2.55727	1.7249	140.0	142.820
8	6	7	1.0882	0.734	140.0	142.820
9	6	8	1.25143	0.8441	70.0	71.414
10	3	11	1.79553	1.2111	140	142.820
11	11	12	2.44845	1.6515	70.0	71.414
12	12	13	2.01317	1.3579	44.1	44.990
13	3	15	1.52348	1.0276	44.1	44.990
14	4	14	2.23081	1.5047	70.0	71.414
15	4	16	1.9702	0.8074	—	—

data sources can be effectively utilized. A key objective of the proposed approach is to extract the most useful information using the fewest possible sensors. By relying on a small set of strategically important buses, the high cost and operational challenges associated with widespread sensor deployment are avoided.

Regarding the selection of input features, an extended set of candidate features was evaluated in the early stages of this work to ensure that the most informative inputs were included. A review of the literature also indicates that most existing studies rely primarily on voltage and current measurements for AI-based schemes. Consistent with these observations, our experiments showed that adding derived features did not improve model performance; in fact, the validation accuracy was consistently higher when only the raw three-phase voltage and current waveforms were used. This suggests that the information contained in the additional features is already embedded within the original waveforms, enabling the model to learn the necessary patterns directly from the time-domain signals. Therefore, it is concluded that using only the voltage and current waveforms provides better performance while keeping the model simpler and more computationally efficient.

The simulation lasted for 0.3 s, with a fault occurring at the 0.1-s mark. Data was recorded continuously during this time (data window) with a sampling period of 0.001 s. This process resulted in a total of 301 data points for each simulation run. The gathered data reflects various states of the network where faults could arise, which were subsequently utilized to train the neural network.

To effectively prepare data for training the neural network, it is essential to explore different scenarios. These scenarios focus on four main aspects of the distribution network: the occurrence of faults in various lines, the types of faults that can happen, variations in fault resistance, and the distance of the fault from the start of the line. By examining these factors, we can create a comprehensive dataset that enhances the neural network's ability to learn and predict fault conditions accurately.

The analysed network, as shown in Figure 1, comprises 15 lines where short-circuit faults can occur. Thus, there are 15 distinct states for this factor. Each of these lines can experience one of ten types of short-circuit faults: three single-phase faults (AG, BG, CG), three line-to-line faults (AB, BC, AC), three line-to-line-to-ground (ABG, BCG, ACG), and one balanced fault (ABC). For training purposes, we focused on faults AG, ABG, and ABC. Therefore, there are three states for this factor.

Fault resistance significantly influences voltage and current levels in the network. To effectively train the neural network, we need to incorporate various fault resistance values. The maximum fault resistance for our analysis is set at 16 Ω based on the network's specifications 7.5 MVA power and an operating voltage of 11 kV. The base impedance is calculated using Formula (1).

$$Z_b = \frac{V_b^2}{S_b} (= 16.1\ \Omega) \tag{1}$$

It is worth noting that this study does not address true high-impedance faults (HIFs). In the technical literature, an HIF is defined as a highly nonlinear phenomenon that typically involves large fault resistance and produces only small, irregular current deviations. Instead, the objective of this study is to evaluate the model under fault resistance values that are higher than those commonly used in related AI-based fault location research, while remaining within a realistic range relative to the system's base impedance. The upper limit of 16 Ω corresponds to approximately

TABLE 3 | Scenarios used to generate data for training the neural network.

Item	Description
Faulty line	15 lines of test system
Type of fault	(A, B, C)-G, (A-B, B-C, C-A), (A-B, B-C, C-A) -G, and A-B-C
Fault resistance	0, 2, 4, 6, 8, 10, 12, 14 and 16 Ω
Fault distance	10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% of the faulted line

1 p.u. of the base impedance for the 11 kV, 7.5 MVA network under study. Previous studies on AI-based fault location typically consider fault resistance-to-base resistance ratios below 0.6 p.u.; in contrast, our study adopts a maximum resistance of 16 Ω, corresponding to 1 p.u. to create more challenging conditions than those examined in earlier works. The fault resistance values used for training consist of 0, 2, 4, 6, 8, 10, 12, 14, and 16 Ω, yielding nine discrete states for this parameter.

As mentioned in Section 1, short-circuit faults can occur across any of the 15 lines in various locations: at the beginning or end of a line or at specific percentages such as 10%, 25% and 30%, or more along the line's length. While theoretically infinite intervals exist for potential faults along a line, this study assumes that faults will occur at intervals of 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% and 90% of the line's initial length, resulting in nine states for this factor as well.

As mentioned in Table 3, considering all four factors together allows for numerous combinations where faults can emerge. Specifically, these factors can generate a total of 3654 unique states, each representing different fault scenarios in the simulated network. During these simulations, we recorded three-phase voltages and current waveforms from both the substation bus and buses containing DG throughout the specified simulation timeframe while training our neural network effectively.

3 | Proposed Fault Location Method

3.1 | Problem Definition and Notation

Consider a three-phase active distribution network with N buses and L distribution lines, indexed by $\{1, 2, \ldots, L\}$, each line potentially hosting multiple fault points. Let $\mathcal{F}$ be the set of all considered fault scenarios, defined by the tuple (ℓ, d, r, ϕ), where $l \in \{1, 2, \ldots, L\}$ denotes the faulted line index, $d \in \{d_1, d_2, \ldots, d_{C_{\text{dist}}}\}$ represents a discretized distance-to-fault value, expressed as a percentage of the line length and enumerated into C_{dist} classes, $r \in \{r_1, r_2, \ldots, r_R\}$ indicates the fault resistance values considered, and $\phi \in \{\text{AG}, \text{ABG}, \text{ABC}, \ldots\}$ denotes the fault type. At selected measurement buses (e.g., the substation and buses equipped with DG units), we measure three-phase voltage and current waveforms over a time window $[0, T']$ seconds. Let the measurement sampling period be Δt, yielding a time series of length $T = \frac{T'}{\Delta t}$ samples. For each fault scenario $s \in \mathcal{F}$, the measured voltage and current at the selected buses are concatenated into a time series $X \in \mathbb{R}^{T \times D}$, where D is the dimension of the combined feature vector at each time step (e.g., three-phase voltages and currents from multiple buses).

In this work, we aim to design a parametric model that, given the measurements X, identifies the faulted line and estimates the distance class along that line. Let θ be the set of learnable parameters (weights and biases, etc.) of this model. We introduce two parametric classification functions as follows:

$$g_{\text{line}}(X; \theta) : \mathbb{R}^{T \times D} \to \{1, 2, \ldots, L\} \tag{2}$$

and

$$g_{\text{dist}}(X; \theta) : \mathbb{R}^{T \times D} \to \{1, 2, \ldots, C_{\text{dist}}\} \tag{3}$$

These functions map the input measurements X to a predicted line index (from the discrete set $\{1, 2, \ldots, L\}$) and a predicted fault distance class $\{1, 2, \ldots, C_{\text{dist}}\}$, respectively. While one approach is to first classify the line and then, conditionally, classify the distance on that line, we here adopt a multi-task learning approach that jointly learns both tasks. By sharing low-level features and intermediate representations, the model captures richer patterns in the input signals, ultimately improving accuracy and robustness under various fault scenarios, fault resistances and network conditions.

3.2 | One-Dimensional Convolutional Neural Networks With Residual Connections

In Section 3.1, we defined a multi-task learning framework that, given time-series measurements $X \in \mathbb{R}^{T \times D}$, aims to simultaneously classify the faulted line and the corresponding distance class. To accomplish this, we require a feature extraction mechanism capable of transforming the raw voltage and current measurements into discriminative representations. 1-D CNNs provide a powerful parametric approach for this task, enabling the model, parameterized by θ, to learn complex temporal patterns inherent in the data. A 1-D CNN operates by convolving filters along the temporal dimension of the input sequence. As illustrated in Figure 2, consider a convolutional filter of size $m = 3$. At a given time step t, the filter is applied to the inputs (x_t, x_{t+1}, x_{t+2}), executing element-wise multiplications with the filter weight (w_1, w_2, w_3) and aggregating the outcomes. The result of this convolutional process is subsequently processed through a nonlinear activation function $(\cdot)$, yielding a feature map value z_t. By sliding the filter across all time steps, we obtain a feature map that encodes local temporal dependencies.

Formally, let $\mathbf{H}^{(0)} = X \in \mathbb{R}^{T \times D_{\text{in}}}$ be the input at the first layer. Consider a convolutional layer parameterized by $\Theta^{(c)} = \{\mathbf{W}^{(c)}, \mathbf{b}^{(c)}\}$ where $\mathbf{W}^{(c)} \in \mathbb{R}^{m \times D_{\text{in}} \times D_{\text{out}}}$ and $\mathbf{b}^{(c)} \in \mathbb{R}^{D_{\text{out}}}$ are the weights and biases associated with the convolutional layer. With zero-padding of the width $m - 1$ at the sequence boundaries to preserve length T, the output $\mathbf{H}^{(1)} \in \mathbb{R}^{T \times D_{\text{out}}}$ is computed as follows:

$$H_{t,o}^{(1)} = f\left(b_o^{(c)} + \sum_{i=1}^{D_{\text{in}}} \sum_{\tau=0}^{m-1} W_{\tau,i,o}^{(c)} . H_{t-1+\tau,i}^{(0)}\right) \tag{4}$$

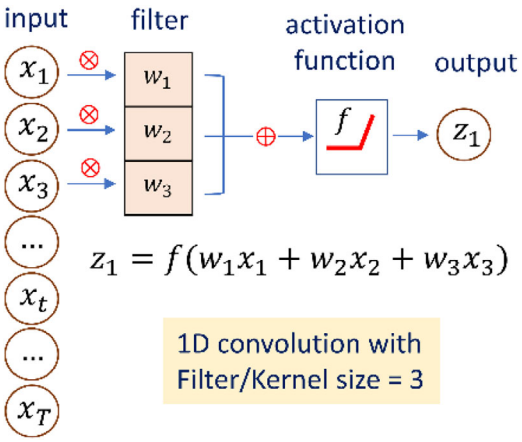

FIGURE 2 | An example of schematic representation of 1-D convolutional operation with a filter/kernel of size 3. The filter slides across the input data, performing element-wise multiplication followed by summation at each position and the resulting output is passed through an activation function to generate a feature map, capturing local temporal patterns critical for time-series analysis.

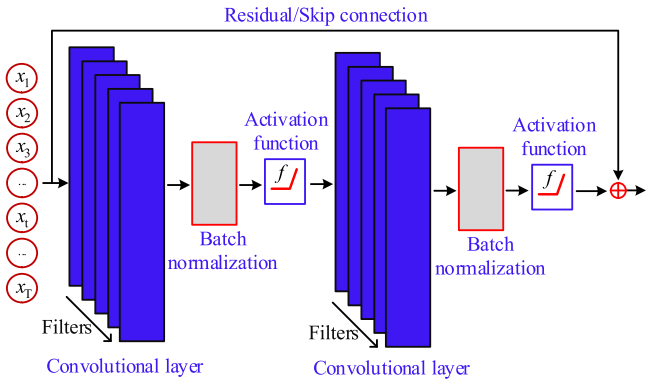

FIGURE 3 | The architecture of a residual block used in this work, incorporating convolutional layers with batch normalization followed by an activation function. The skip connection enables the input to bypass intermediate layers, directly connecting to the output to enhance gradient flow, mitigating vanishing gradient issues, and improving model convergence.

for $t = 1, \ldots, T$ and $o = 1, \ldots, D_{\text{out}}$. Stacking multiple convolutional layers refines the extracted features, enabling the model to capture more intricate temporal correlations related to different fault types, resistances, and operating conditions defined in Section 3.1. However, as the network depth increases, direct stacking of convolutional layers may lead to training difficulties such as vanishing or exploding gradients. To address this, we incorporate residual (skip) connections, as shown in Figure 3. A residual block consists of two convolutional layers with batch normalization (that normalize the distribution of data into the layer) and nonlinear activations, and a skip connection that adds the block input directly to its output. This skip connection ensures that the gradient can propagate directly through the identity path, simplifying the training of deeper architectures and allowing the model to learn richer representations. Additionally, to further enhance the learned representations and reduce computational complexity, pooling layers are employed. Pooling operations (e.g., max pooling or global average pooling) condense the temporal dimension, focusing on the most salient features. By retaining essential characteristics and discarding redundant information, pooling facilitates robust feature extraction from long sequences of voltage and current measurements.

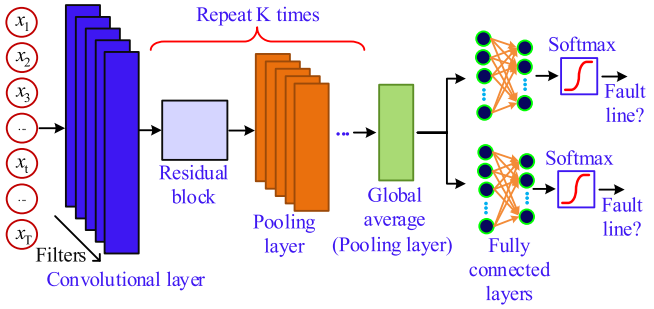

FIGURE 4 | Our proposed model that integrates a sequence of convolutional layers followed by residual blocks to extract robust features from voltage and current measurements for fault detection and localization. This model uses two separate fully connected output layers with softmax activation: one to classify the faulty line and the other to estimate the fault distance, and is trained by minimizing the cross-entropy loss function between the softmax outputs and the ground truth labels.

Following feature extraction, the model employs fully connected (FC) layers to map the learned representations onto discrete classes. As illustrated in Figure 4, we adopt a multi-task learning strategy by using two separate FC layers and softmax outputs: one head for classifying the faulted line, and another for classifying the fault distance. Let $\mathbf{W}^{(\text{line})} \in \mathbb{R}^{D_{\text{out-final}} \times L}$ and $\mathbf{W}^{(\text{dist})} \in \mathbb{R}^{D_{\text{out-final}} \times C_{\text{dist}}}$ be the weights of the final classification layers for line and distance classification, respectively, where L is the number of lines and C_{dist} is the number of distance classes defined in Section 3.1. For the line classification, we have:

$$\hat{y}_i^{(\text{line})} = \frac{\exp\left(w_i^{(\text{line})} \times h\right)}{\sum_{j=1}^{L} \exp\left(w_j^{(\text{line})} \times h\right)} \; ; \; i = 1, \ldots, L \tag{5}$$

and for the distance classification:

$$\hat{y}_j^{(\text{dist})} = \frac{\exp\left(w_j^{(\text{dist})} \times h\right)}{\sum_{k=1}^{C_{\text{dist}}} \exp\left(w_k^{(\text{dist})} \times h\right)} \; ; \; j = 1, \ldots, C_{\text{dist}} \tag{6}$$

where $w_i^{(\text{line})}$ and $w_j^{(\text{dist})}$ are the ith and jth column vectors of $\mathbf{W}^{(\text{line})}$ and $\mathbf{W}^{(\text{dist})}$, respectively. These softmax outputs represent the model's predicted probability distributions over the possible line and distance classes.

Training this architecture involves minimizing a combined loss function that accounts for both tasks. Let $y^{(\text{line})}$ and $y^{(\text{dist})}$ be the one-hot encoded ground-truth labels for line and distance classification, respectively. We can define the cross-entropy losses as follows:

$$\mathcal{L}_{\text{line}} = -\sum_{i=1}^{L} y_i^{(\text{line})} \log\left(\hat{y}_i^{(\text{line})}\right) \tag{7}$$

$$\mathcal{L}_{\text{dist}} = -\sum_{j=1}^{C_{\text{dist}}} y_j^{(\text{dist})} \log\left(\hat{y}_j^{(\text{dist})}\right) \tag{8}$$

Therefore, the total loss for training the model presented in Figure 4 would be,

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{line}} + \mathcal{L}_{\text{dist}} \tag{9}$$

By simultaneously optimizing both line and distance classification tasks through minimizing $\mathcal{L}_{\text{total}}$, the model learns shared representations that improve overall fault localization performance and enhance generalization to diverse operational conditions.

3.3 | A Discussion on the Technical Basis and Performance Advantages of the Proposed Method

In this section, a detailed explanation is provided regarding how the proposed method addresses the challenges associated with distribution networks characterized by high branching, limited observability, varying fault resistance, and the presence of IBDGs. Although the technical components of the proposed scheme are defined in the previous subsections, the conceptual basis that explains its effectiveness under the aforementioned operating conditions is elaborated here.

3.3.1 | 3.3.1 Simulation Strategy Designed for Challenging Operating Conditions

Unlike many existing works that rely on simplified test systems, the simulation framework employed in this study is designed to explicitly represent conditions under which conventional methods tend to perform poorly.

a. Presence of IBDGs:

Three IBDGs are installed at different locations in the network. During faults, IBDGs respond with fast control actions and limit their output current, which creates transient behaviours that are quite different from those in systems with traditional synchronous machines. These characteristics are fully included in the generated dataset, allowing the trained models to learn fault behaviours that are often missed in conventional studies.

b. Fault resistance variation:

Fault resistance values up to 16 Ω (corresponding to the base impedance of the system) are included. Although these values do not correspond to HIFs, they cover the practical range of distribution-level faults and introduce cases where fault signatures are significantly damped.

c. Limited measurement availability:

Only the substation and DG buses are used as measurement points. This constraint requires the model to infer both faulted sections and distances for branches lacking direct instrumentation, thereby reflecting realistic operational conditions.

d. High network branching:

The test system contains numerous radial branches and laterals, increasing the difficulty of distinguishing between faults occurring on sections with similar electrical characteristics.

All comparative methods are evaluated under identical conditions to ensure a consistent and equitable performance assessment.

3.3.2 | Architectural Features Addressing These Challenges

To accommodate the complexities embedded in the simulated dataset, the proposed model incorporates two architectural elements designed to enhance robustness and generalization.

a. Residual connections for weak and nonlinear transients:

The fault resistance produces low-energy and heavily damped transient waveforms. In addition, nonlinear dynamics introduced by IBDG control loops further distort these signatures. In standard deep neural networks, such subtle features often diminish due to vanishing gradients during training. By employing residual (skip) connections, gradient propagation is preserved across deeper layers. This enables the network to learn high-order temporal correlations and enhances its capacity to detect weak disturbances and nonlinear transient patterns associated with Rf variation and IBDG behaviours.

b. Multi-task learning as implicit regularization:

The model is trained to perform both fault-line identification and fault-distance estimation using a shared feature extractor. This multi-task learning structure imposes a regularizing effect by encouraging the extraction of feature representations that remain meaningful across different tasks. As a result, overfitting to noise or scenario-specific artifacts is reduced, which is particularly beneficial when measurement availability is limited.

4 | Simulation Results

4.1 | Data Simulation and Model Structures

In this work, we use the simulated dataset described in Section 2, which comprises 3654 time-series samples, each spanning 300 ms of measurements and containing 16 features. The dataset is divided into training and testing subsets at an 80–20 split, and 20% of the training portion is reserved as a validation set for hyperparameter tuning. Details regarding the model architectures, including layer dimensions and number of parameters, are provided in Table 4.

The 80% (training + validation)/20% (test) split was applied to a synthetic dataset of 3654 fault scenarios specifically designed

TABLE 4 | A detailed presentation of the models' architectures used in this work for CNN model 1, CNN model 2, LSTM model, and the proposed model. All models are trained using Adam optimization with a learning rate of 0.001. The models were trained with a batch size of 32 for a maximum of 400 epochs.

CNN model 1 (inspired by [9])

Model complexity:
x = Conv1D (filters = 6, kernel_size = 9) (In)
x = AveragePooling1D(pool_size = 2) (x)
x = Conv1D (filters = 12, kernel_size = 9) (x)
x = AveragePooling1D(pool_size = 2) (x)
x = Flatten () (x)
x = Dense (168) (x)
o1 = Dense (15, activation = 'softmax') (x)
o2 = Dense (9, activation = 'softmax') (x)

CNN model 2 (inspired by [8])

Model complexity:
x = Conv1D (filters = 32, kernel_size = 4) (In)
x = AveragePooling1D (pool_size = 3) (x)
x = Conv1D (filters = 48, kernel_size = 6) (x)
x = AveragePooling1D (pool_size = 3) (x)
x1 = Flatten () (x)
o1 = Dense (15, activation = 'softmax') (x1)
x2 = Conv1D (filters = 64, kernel_size = 4) (x)
x2 = Flatten () (x2)
x2 = Dense (64) (x2)
o2 = Dense (9, activation = 'softmax') (x2)

LSTM model (inspired by LSTM module in [10])

Model complexity:
x = LSTM (units = 256, return_sequences = True) (In)
x = Dropout (0.5) (x)
x = LSTM (256, return_sequences = True) (x)
x = Dropout (0.5) (x)
x = Flatten () (x)
o1 = Dense (15, activation = 'softmax') (x)
o2 = Dense (9, activation = 'softmax') (x)

Our model (1-D CNN + Residual blocks)

x = ResBlock (filters = 128)(x)
3 times : {x = MaxPooling1D (pool_size = 2)(x)
x = Dropout (0.1)(x)
x = ResBlock (filters = 768)(x)
2 times : {x = MaxPooling1D (pool_size = 2)(x)
x = Dropout (0.1)(x)
x = ResBlock (filters = 512) (x)
x = GlobalAveragePooling1D ()(x)
o1 = Dense (15, activation = 'softmax') (x)
o2 = Dense (9, activation = 'softmax') (x)
where each ResBlock (filters = K) (In) is:
x = Conv1D (filters = K, kernel_size = 3) (In)
x = BatchNormalization() (x)
x = Activation('relu') (x)
x = Conv1D (filters = K, kernel_size = 3) (x)
x = BatchNormalization() (x)
x = Activation('relu') (x)

to maximize variability and cover a large range of conditions. In particular,

- Uniform sampling: Faults were sampled at 100 equally spaced locations along every line, covering 10 fault types, and across a wide range of fault resistance values (up to 16 Ω) and fault inception angles.
- Unseen data: The 20% test set contained fault locations, resistances and inception angles that were completely unseen during the training process, ensuring that the test set truly measured generalization rather than memorization.

To rigorously confirm that the performance metrics obtained from this initial random 80/20 split were not an artifact of that particular data partitioning, we also performed a comprehensive five-fold cross-validation (CV) during our model development. The average performance metrics (balanced accuracy, F1-score) obtained from the five-fold cross-validation process were highly consistent with the initial results reported in the manuscript (Table 5). The standard deviation across the five folds was minimal, confirming the stability of the performance. The consistency between the initial 80/20 test set result and the average five-fold CV result strongly validates that the initial random partitioning of our comprehensive, high-variability synthetic dataset was sufficient for accurately estimating the model's true generalization capability.

In this work, both faulted line identification and fault distance estimation are framed as multi-class classification problems, where each sample is assigned to one of $L = 15$ possible line classes or one of $C_{\text{dist}} = 9$ possible distance classes. To assess the effectiveness of the suggested models, let TP_i, TN_i, FP_i and FN_i denote, respectively, the true-positives, true-negatives, false-positives and false-negatives associated with class i. The most common evaluation metric is the overall accuracy, given by:

$$\text{Accuracy} = \sum_{i=1}^{C} \frac{\text{TP}_i + \text{TN}_i}{\text{TP}_i + \text{TN}_i + \text{FP}_i + \text{FN}_i} \tag{10}$$

where C is the total number of classes in the considered task (i.e., C = 15 for line classification and C = 9 for distance classification). While accuracy provides an aggregate measure of correctness, it may lead to misinterpretations when class imbalance is present. More precisely, a model that persistently predicts the majority class can attain an inaccurately high accuracy, despite underperforming on minority classes. To address this limitation, we employ balanced accuracy defined as follows:

$$\text{Balanced accuracy} = \frac{1}{C} \sum_{i=1}^{C} \frac{\text{TP}_i}{\text{TP}_i + \text{FN}_i} \tag{11}$$

In addition, other class-sensitive metrics, including precision, recall, and F1-score, are defined for each class as follows [24]:

$$\begin{aligned} \text{Precision} &= \frac{\text{TP}}{\text{TP} + \text{FP}} \\ \text{Recall} &= \frac{\text{TP}}{\text{TP} + \text{FN}} \\ \text{F1-score} &= 2\left(\frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}\right) \end{aligned} \tag{12}$$

TABLE 5 | The results of the proposed scheme based on artificial intelligence.

Model		Balanced accuracy	F1-score	Precision	Recall
CNN model 1 (inspired by [9])	Fault line	88.29	87.80	88.78	87.96
	Fault distance	50.41	49.63	51.96	50.31
CNN model 2 (inspired by [8])	Fault line	86.96	85.75	86.20	87.04
	Fault distance	48.67	42.07	44.51	48.77
LSTM model (inspired by LSTM module in [10])	Fault line	89.78	89.46	89.73	89.51
	Fault distance	90.28	90.08	90.78	90.12
Our model (1-D CNN + residual blocks)	Fault line	93.80	93.59	94.47	93.83
	Fault distance	96.94	96.92	96.94	96.91

When evaluating a multi-class classifier, it is common to aggregate these class-level metrics into a single performance indicator. Since the distribution of classes in the dataset can be highly imbalanced (e.g., certain lines or distance classes occur more frequently), we adopt a weighted averaging strategy based on the total number of samples at each class.

4.2 | Results and Discussion

Based on the results given in Table 5, our proposed method, which incorporates residual blocks, demonstrates exceptional performance in fault localization for inverter-based active distribution networks. Specifically, it achieves an accuracy of approximately 94% in identifying faulty components and 97% in estimating fault distances along the line. In comparison, the CNN model 1 [9], when evaluated under the conditions specified in Section 1, achieves an accuracy of 88% for fault identification and 50% for fault distance estimation. Similarly, the approach using CNN model 2 [8] achieves 87% accuracy for identifying faulty sections and 48% for measuring fault distances. The method based on the LSTM model [10] reports a balanced accuracy of about 90% for both fault identification and fault distance measurement. The superior performance of our method can be attributed to the use of residual blocks within the 1-D CNN architecture. These residual blocks mitigate the vanishing gradient problem, enabling deeper network layers to effectively learn complex features from the input data. By preserving gradient flow during training, our architecture captures complex temporal patterns in voltage and current signals, which are critical for accurately identifying faults line and estimating their distances. Furthermore, the multi-task learning framework employed in our model ensures that shared representations are leveraged to improve performance across both tasks.

These findings highlight that our approach meets the requirements of realistic distribution network conditions, including high fault resistance, limited measurement points, and the presence of inverter-based distributed generators. This makes our method highly suitable for practical deployment in modern power systems, where reliability and precision are important.

Regarding the evaluation metrics, we have deliberately focused on balanced accuracy and F1-score due to the practical importance of correctly identifying all fault types (e.g., AG, ABC, etc.). Although our dataset is large, these classes are not evenly distributed in terms of frequency. Standard accuracy can be highly skewed toward the majority classes, which would lead to a superficial assessment and potentially misleading model comparisons. Balanced accuracy is a critical metric for this multi-class classification problem because it is simply the average of the per-class accuracy. This ensures that the model is penalized equally for misclassifying a rare, but safety-critical, fault type as it is for misclassifying a common one.

The decision to use a discrete classification approach for fault distance estimation (as opposed to continuous regression) was a deliberate, practical choice driven by the need for high-confidence results in a real-world operational environment. A continuous regression model provides a single, floating-point distance (e.g., 2.345 km). This value is inherently associated with a mean absolute error, which can lead to ambiguity for repair crews (e.g., 'The fault is at 2.3 km ± 150 m'). Our classification approach is designed to:

- Reduce ambiguity as it classifies the fault into one of nine specific bins (or sections) per line, which directly correspond to easily identifiable, physically defined sections of the distribution line.
- Ensure actionable information, as this provides crisp, actionable zones for the repair crews, directing them to a definite segment (e.g., 'Section 2 of Line L12–L13') rather than a continuous distance with an associated error margin.

It is worth mentioning that, while our study does not report a specific computational response time, the proposed fault location method is designed as an offline process. In practical engineering applications, protection relays first detect a fault, and an event recorder captures the voltage and current signals immediately before and after the fault. This capability is enabled by the triggering system and the sequential event recording functionality of event recorders, which ensures that all relevant measurements are accurately captured around the fault occurrence. The recorded data is then processed by our model to identify the fault segment and estimate the fault distance. All measurements are synchronized using timestamps to ensure proper alignment for analysis. Although the model is not intended for real-time

protection, this offline approach is suitable for post-event fault analysis and provides accurate fault location information without interfering with system operation.

5 | Conclusion

In conclusion, the proposed fault location approach in this paper is presented through the utilization of a modified 1-D CNN with residual connections. Various types of faults are considered, including line-to-ground (LG), line-to-line (LL), double line-to-ground (LLG) and balanced (LLL) faults, which are essential for accurately assessing the performance of the suggested approach. To evaluate the effectiveness of the suggested scheme, extensive time-domain simulations are conducted using a modified 15-bus IEEE network. This network features multiple branches and includes inverter-based distributed generation sources, which introduce unique challenges for fault detection and localization.

The performance evaluation reveals that our approach achieves an outstanding accuracy of 94% in identifying faulty components and 97% in estimating fault distances along the distribution lines. This high level of accuracy indicates that the proposed approach is highly effective in real-world applications where precise fault detection is critical. This performance represents a notable improvement over existing methods. For instance, CNN Model 1 achieved a balanced accuracy of only 88.29% for fault identification and a mere 50.41% for fault distance estimation. Similarly, CNN Model 2 reported accuracies of 86.96% for identifying faults and just 48.67% for estimating distances. The LSTM model previously reported a balanced accuracy of approximately 89.78% for fault identification and 90.28% for fault distance estimation. In contrast, our model demonstrates superior performance with a balanced accuracy of 93.80% for fault line identification and an impressive 96.94% for fault distance estimation. Additionally, our model achieves F1-scores of 93.59% and 96.92%, respectively, reflecting its ability to balance precision and recall effectively.

These results highlight the robustness of the proposed scheme in addressing the complexities associated with modern distribution networks characterized by high fault resistance, limited measurement points, multilateral systems, and different types of inverter-based resources.

Finally, it is worth to note that the proposed model's structural advantages are not merely empirical, but are rooted in specific deep learning principles designed to overcome the unique challenges of modern distribution networks. These advantages can be summarized into three theoretical pillars:

1. Resilience through residual connections (ResNet): The core structural advantage is the incorporation of residual blocks (ResNet architecture). This design directly addresses the difficulty of training deep neural networks caused by the vanishing gradient problem.
   - Addressing signal damping: The fault resistance dampens the fault signal, resulting in weak (low-magnitude) feature gradients. In a standard deep CNN, these subtle gradients often vanish before reaching the initial layers during backpropagation, preventing the model from learning the characteristics of faults with resistance.
   - The structural solution: The skip connections allow the gradient to bypass one or more layers and flow directly to earlier layers. This ensures that even the subtle feature information required to detect dampened signals is preserved and learned effectively, enabling us to build a deeper network that is necessary for accurate fault location.
2. Adaptive feature extraction with 1-D CNNs: The choice of a 1-D CNN structure provides a theoretical advantage by abstracting the need for manual feature engineering.
   - Handling IBDG dynamics: Unlike traditional methods that rely on symmetrical components or pre-defined features, the 1-D convolutional filters automatically learn the optimal temporal correlations from the raw, three-phase voltage and current waveforms. This allows the model to inherently capture the highly complex and non-linear and dynamic transient signatures caused by IBDG control strategies during faults. The network is not limited by human-designed feature sets; it learns the distinguishing features autonomously.
3. Generalization via multi-task learning: Our model is structured as a multi-task learning model, which shares a single residual backbone for two distinct outputs—fault line classification and fault distance classification (binned).
   - The regularization effect: The shared backbone acts as a powerful structural regularizer. By forcing the model to learn intermediate feature representations that must be simultaneously useful for both tasks, the model is prevented from overfitting to task-specific noise. This leads to the extraction of more robust, generalized features that enhance the model's performance on unseen fault data (test set) and challenging conditions (limited meters).

Author Contributions

Morteza Behbahanipour: conceptualization (equal), data curation (equal), formal analysis (equal), investigation (equal), resources (equal), software (equal), validation (equal), writing – original draft (equal), writing – review and editing (equal). **Seyed Fariborz Zarei:** conceptualization (equal), formal analysis (equal), investigation (equal), methodology (equal), project administration (equal), resources (equal), supervision (equal), validation (equal), writing – original draft (equal), writing – review and editing (equal). **Mohammadhadi Shateri:** conceptualization (equal), data curation (equal), formal analysis (equal), methodology (equal), resources (equal), supervision (equal), validation (equal), writing – original draft (equal), writing – review and editing (equal).

Acknowledgements

The authors acknowledge the use of Perplexity AI (https://www.perplexity.ai) and ChatGPT (OpenAI) to enhance the clarity and language of the manuscript. The tool was used solely for checking sentence grammar and writing during the drafting phase. All suggestions and generated content were critically reviewed and edited by the authors to ensure compliance with academic standards.

Funding

The authors have nothing to report.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

1. M. Yadegar, S. F. Zarei, N. Meskin, and F. Blaabjerg, "A Distributed High-Impedance Fault Detection and Protection Scheme in DC Microgrids," *IEEE Transactions on Power Delivery* 39, no. 1 (2023): 141–154.

2. P. Stefanidou-Voziki, N. Sapountzoglou, B. Raison, and J. L. Dominguez Garcia, "A Review of Fault Location and Classification Methods in Distribution Grids," *Electric Power Systems Research* 209 (2022): 108031, https://doi.org/10.1016/j.epsr.2022.108031.

3. S. F. Zarei, H. Mokhtari, and F. Blaabjerg, "Fault Detection and Protection Strategy for Islanded Inverter-Based Microgrids," *IEEE Journal of Emerging and Selected Topics in Power Electronics* 9, no. 1 (2019): 472–484, https://doi.org/10.1109/JESTPE.2019.2962245.

4. S. F. Zarei, M. A. Ghasemi, and S. Khankalantary, "Current Limiting Strategy for Grid-Connected Inverters Under Asymmetrical Short Circuit Faults," *International Journal of Electrical Power & Energy Systems* 131 (2021): 107020, https://doi.org/10.1016/j.ijepes.2021.107020.

5. S. Gururajapathy, H. Mokhlis, and H. Illias, "Fault Location and Detection Techniques in Power Distribution Systems With Distributed Generation: A Review," *Renewable and Sustainable Energy Reviews* 74 (2017): 949–958, https://doi.org/10.1016/j.rser.2017.03.021.

6. B. P. Morteza and Z. S. Fariborz, "A Fault Location Scheme for Distribution Networks with Inverter-Based Resources," in *Proceedings of the Power Electronics and Drives: Systems and Technologies Conference (PEDSTC 2024)* (Scientific Information Database, 2024), 1132495, https://sid.ir/paper/1132495/en.

7. P. Stefanidou Voziki, N. Sapountzoglou, B. Raison, and J. L. Dominguez-Garcia, "A Review of Fault Location and Classification Methods in Distribution Grids," *Electric Power Systems Research* 209 (2022): 108031, https://doi.org/10.1016/j.epsr.2022.108031.

8. Y. Yu, M. Li, T. Ji, and Q. H. Wu, "Fault Location in Distribution System Using Convolutional Neural Network Based on Domain Transformation," *CSEE Journal of Power and Energy Systems* 7, no. 3 (2021): 472–484, https://doi.org/10.17775/CSEEJPES.2020.01620.

9. M.-F. Guo, J.-H. Gao, X. Shao, and D.-Y. Chen, "Location of Single-Line-to-Ground Fault Using 1-D Convolutional Neural Network and Waveform Concatenation in Resonant Grounding Distribution Systems," *IEEE Transactions on Instrumentation and Measurement* 70 (2020): 1–9, https://doi.org/10.1109/TIM.2020.3014006.

10. J. Yuan and Z. Jiao, "Faulty Feeder Detection for Single Phase-to-ground Faults in Distribution Networks Based on Patch-to-Patch CNN and Feeder-to-Feeder LSTM," *International Journal of Electrical Power & Energy Systems* 147 (2023): 108909, ISSN 0142–0615, https://doi.org/10.1016/j.ijepes.2022.108909.

11. S. Jamali, A. Bahmanyar, and S. Ranjbar, "Hybrid Classifier for Fault Location in Active Distribution Networks," *Protection and Control of Modern Power Systems* 5, no. 2 (2020): 1–9.

12. B. Nguyen, T. Vu, T.-T. Nguyen, M. Panwar, and R. Hovsapian, "Spatial-Temporal Recurrent Graph Neural Networks for Fault Diagnostics in Power Distribution Systems," *IEEE Access* 11 (2023): 46039–46050, https://doi.org/10.1109/ACCESS.2023.3273292.

13. D. Chanda and N. Y. Soltani, "A Heterogeneous Graph-Based Multi-Task Learning for Fault Event Diagnosis in Smart Grid," preprint, arXiv, September 18, 2023, https://arxiv.org/abs/2309.09921.

14. J. Hu, W. Hu, J. Chen, et al., "Fault Location and Classification for Distribution Systems Based on Deep Graph Learning Methods," *Journal of Modern Power Systems and Clean Energy* 11, no. 1 (2023): 35–51, https://doi.org/10.35833/MPCE.2022.000204.

15. F. Mumtaz, K. Imran, H. Rehman, and H. Qureshi, "Hardware Supported Fault Detection and Localization Method for AC Microgrids Using Mathematical Morphology With State Observer Algorithm," *IEEE Access* 12 (2024): 12446–12457, https://doi.org/10.1109/ACCESS.2024.3354790.

16. Y. Chen, Q. Wang, X. Cai, and H. Jiang, "Operation Status Identification and Fault Location of Power System Based on SVM," in *Proceedings of the 2023 6th International Conference on Energy, Electrical and Power Engineering (CEEPE)* (IEEE, 2023), 749–753, https://doi.org/10.1109/CEEPE58418.2023.10165902.

17. A. Srivastava and S. Parida, "A Robust Fault Detection and Location Prediction Module Using Support Vector Machine and Gaussian Process Regression for AC Microgrid," *IEEE Transactions on Industry Applications* 58, no. 1 (2021): 930–939, https://doi.org/10.1109/TIA.2021.3129982.

18. M. Li, H. Zhang, T. Ji, and Q. H. Wu, "Fault Identification in Power Network Based on Deep Reinforcement Learning," *CSEE Journal of Power and Energy Systems* 8, no. 3 (2022): 721–731, https://doi.org/10.17775/CSEEJPES.2020.04520.

19. X. Wang, X. Wei, J. Gao, G. Song, M. Kheshti, and L. Guo, "High-Impedance Fault Detection Method Based on Stochastic Resonance for a Distribution Network With Strong Background Noise," *IEEE Transactions on Power Delivery* 37 (2022): 1004–1016, https://doi.org/10.1109/TPWRD.2021.3075472.

20. N. Sapountzoglou, J. Lago, and B. Raison, "Fault Diagnosis in Low Voltage Smart Distribution Grids Using Gradient Boosting Trees," *Electric Power Systems Research* 182 (2020): 106254, https://doi.org/10.1016/j.epsr.2020.106254.

21. A. Gopi and P. A. D. V. Raj, "Distributed Generation for Line Loss Reduction in Radial Distribution System," in *Proceedings of the International Conference on Emerging Trends in Electrical Engineering and Energy Management* (IEEE, 2012), 29–32.

22. S. F. Zarei and M. Parniani, "A Comprehensive Digital Protection Scheme for Low-Voltage Microgrids With Inverter-Based and Conventional Distributed Generations," *IEEE Transactions on Power Delivery* 32, no. 1 (2016): 441–452, https://doi.org/10.1109/TPWRD.2016.2566264.

23. M. Behbahanipour, S. F. Zarei, and M. Shateri, "An Impedance-Based Approach for Locating Short-Circuit Faults in Inverter-Based Active Distribution Networks," *International Journal of Industrial Electronics Control and Optimization* 7, no. 3 (2024): 225–233.

24. T. Saito and M. Rehmsmeier, "The Precision-Recall Plot is More Informative Than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets," *PLoS One* 10, no. 3 (2015): e0118432, https://doi.org/10.1371/journal.pone.0118432.