EAI Endorsed Transactions
on Energy Web
Research Article EAI.EU

A Dual-Graph Physics-Informed Graph Attention Network for Fault Location in Active Distribution Networks

Yi Zheng[1], Ruiqiang Zhang[1], Pengfei Qu[1], Ying Zhang[1], Zhongyou Shen[2] and Yang Lu[3*]

[1] State Grid Sichuan Information and Communication Company, Chengdu 610000, China
[2] State Grid Luzhou Power Supply Company, Luzhou 646000, China
[3] Sichuan Energy Internet Research Institute, Tsinghua University, Chengdu 610000, China

Abstract

INTRODUCTION: Fault localization in active distribution networks (ADNs) is challenging because feeder topology and electrical coupling may become inconsistent under operating conditions with high penetration of distributed energy resources. Moreover, purely data-driven models often provide limited physical interpretability.
OBJECTIVES: This study aims to develop a dual-graph learning framework for fault localization in ADNs, with particular attention to scenarios where topology-based message passing alone may be insufficient due to topology–electrical mismatch.
METHODS: A dual-graph physics-informed graph attention network (DG-PIGAT) is proposed. The method uses a topology graph to describe feeder connectivity and an electrical-distance graph to characterize impedance-related coupling among buses. Electrical-distance-guided propagation is introduced to improve feature aggregation under electrical coupling variations, while Kirchhoff's Current Law (KCL)-related residual information is incorporated as a weak physical regularization term to support physically plausible learning.
RESULTS: Experiments on the IEEE 33-bus benchmark under four consolidated scenarios show that DG-PIGAT achieves the highest mean accuracy across the tested scenarios. Its clearest advantage appears under topology–electrical mismatch, where it achieves 86.92% accuracy and outperforms GraphSAGE by 8.29 percentage points (p = 0.0025). Sensitivity and robustness tests also indicate better tolerance to measurement noise, missing data, and topology–electrical inconsistency.
CONCLUSION: The results suggest that combining feeder topology with impedance-related electrical coupling is beneficial for robust fault localization in ADNs. The KCL-related component should be interpreted conservatively as a physical regularization mechanism rather than the dominant source of accuracy improvement.

Keywords: active distribution networks; fault location; graph neural networks; deep learning; physics-informed machine learning; intelligent fault diagnosis; dual-graph modeling; distributed energy resources.

Received on 25 January 2026, accepted on 15 April 2026, published on 22 September 2026

doi: 10.4108/ew.15083

1. Introduction

The accelerated deployment of distributed energy resources (DERs) is fundamentally reshaping traditional distribution networks into active distribution networks (ADNs), where bidirectional power flows and dynamic topology changes become common. This transition introduces increased operational uncertainty and makes fault location more challenging [1]. Accurate fault location, as a key component

*Corresponding author. Email: luyang@tsinghua-eiri.org

of self-healing control, is essential for maintaining power supply reliability [2].

Conventional paradigms, encompassing both traveling-wave and impedance-based approaches, are strictly contingent upon precise system parameters and high-quality data. Their efficacy often diminishes under the intricate operating states of modern ADNs [3,4].

Recently, graph neural networks (GNNs) have emerged as a potent tool for fault diagnosis, owing to their inherent capacity to handle the non-Euclidean nature of power grids. Nevertheless, existing GNN-based approaches still exhibit several limitations when applied to ADNs. On the one hand, the mismatch between physical topology and electrical characteristics prevents the model from accurately capturing the true coupling relationships among nodes. On the other hand, purely data-driven models lack physical inductive bias, which may lead to results that violate fundamental electrical laws. In addition, the fusion of multi-dimensional features is often insufficiently coordinated, further constraining model performance.

To address these gaps, this paper introduces DG-PIGAT, a framework that shifts the paradigm from simple topology-based message passing to a dual-graph physics-informed architecture. Specifically, we represent the distribution network through coupled static topology and electrical distance graphs, an approach that preserves informative coupling even when physical connectivity is partially misaligned. To ensure physical consistency, the learning process is guided by KCL-related residual information and weak physical regularization. The practical value of this design is validated on the IEEE 33-bus system, where DG-PIGAT demonstrates not only competitive accuracy in standard scenarios but also a statistically significant advantage in handling topology-electrical mismatches—a notorious challenge for conventional GNNs.

2. Related Work

Existing studies on distribution-network fault localization can be organized into three representative research streams: data-driven models, graph neural network (GNN)-based models, and physics-informed approaches. The following review summarizes these directions and clarifies the position of the proposed method.

2.1. Data-Driven Fault Location Methods

As operational measurements have become more accessible, data-driven techniques have emerged as an important option for distribution-network fault localization [5]. Earlier studies commonly adopted convolutional neural network (CNN) and recurrent neural network (RNN) architectures to learn temporal characteristics from voltage and current waveforms, achieving satisfactory results when the feeder topology was unchanged [6,7].

A common limitation of such models is that they usually process node measurements as independent sequences in Euclidean space. Under this setting, the structural relationships among nodes are not explicitly represented, and the spatial propagation of fault-related information is therefore difficult to capture. This issue becomes more pronounced when the measurement layout is sparse or when the topology changes during operation.

2.2. Graph Neural Network-Based Methods

Because distribution networks are naturally graph-structured systems, GNNs have gradually been introduced into fault analysis tasks. By passing messages among neighboring nodes, these models can describe network interactions more directly than conventional sequence-based methods. Graph convolutional networks provide a representative message-passing formulation for semi-supervised graph learning [8], and deep graph convolutional networks have also been used for fault location in power distribution systems [9]. Among graph-based models, graph attention networks (GATs) are particularly attractive because they assign adaptive weights to neighboring nodes during feature aggregation [10]. In addition, GraphSAGE, GATv2, and ChebNet provide representative inductive aggregation, dynamic graph attention, and spectral graph convolution baselines, respectively [11–13].

Recent studies have applied GNN-based models to fault detection and fault location in distribution networks. Existing work has considered weak fault feature perception, improved generalization, cost-sensitive learning, and multi-scale feature extraction [14–16]. Graph attention mechanisms have also been incorporated into fault location models, where they help capture heterogeneous dependencies among network nodes [17].

Some studies have further extended this line of work by introducing more than one graph structure. KG-FLoc, for example, combines equipment topology with knowledge graph information to improve fault localization in secondary circuits [18]. GNN-based models have also been explored in broader power system tasks, suggesting that graph-based learning is effective for complex interconnected systems [19].

Even so, most current GNN-based methods still construct graph structures mainly from physical connectivity. Differences in electrical coupling strength are often not represented explicitly, which may reduce performance under long-feeder, high-impedance, or weak-fault conditions.

2.3. Physics-Informed Methods

Physics-informed machine learning embeds physical laws or equation residuals into neural-network training, with PINNs providing a representative framework for forward and inverse problems [20,21]. In power systems, physics-aware and physics-informed neural networks have been studied for distribution-system state estimation and other learning tasks [22,23]. This idea has motivated several recent studies that combine physical knowledge with graph neural networks. Typical examples include graph attention models with embedded physical constraints and physics-preserving graph networks [24–26]. These studies suggest that introducing physical inductive bias can improve out-of-distribution generalization [27].

Despite these advances, physical information in most existing methods is still imposed mainly through the loss function. In practice, this means that physical laws are used primarily to regulate the model output, while the feature propagation process itself remains weakly constrained. As a result, the improvement in physical consistency is often limited.

2.4. Summary of Existing Methods

Based on the above review, existing methods still face several limitations in addressing fault location problems in active distribution networks:

1. Existing GNN-based models do not fully capture the discrepancy between topological connectivity and electrical coupling;
2. Physical information is typically introduced as a post hoc constraint, with limited influence on the feature propagation process;
3. Multi-graph approaches mainly focus on data fusion, while collaborative modeling aligned with power system physical mechanisms remains insufficiently explored.

Motivated by the above gaps, this study develops a Dual-Graph Architecture Phys-ics-Informed Graph Attention Network (DG-PIGAT).

3. Methodology

The overall methodology of the proposed DG-PIGAT model is presented in this section. First, the problem formulation and the overall framework are introduced. Then, the dual-graph construction process is described, including the static topology graph and the dynamic electrical distance graph. Subsequently, a dual-branch feature extraction mechanism based on graph attention networks is developed. Finally, the feature fusion strategy and the physics-informed loss function are presented.

3.1. Problem Formulation and Overall Framework

The distribution network is modeled by a graph $\boldsymbol{G} = (V, E)$, whose node set $V = \{v_1, \dots, v_N\}$ encompasses buses, loads, or distributed energy resource (DER) connection points, and E represents the set of distribution lines. Each edge $(i, j) \in E$ is associated with physical parameters, including resistance r_{ij}, reactance χ_{ij}, and impedance magnitude z_{ij}, which are used to construct the node impedance matrix $\boldsymbol{Z}$ for subsequent electrical distance calculation.

For each node $i \in V$, the measurement feature vector is defined as:

$$\boldsymbol{x}_i = [U_i,\ P_i,\ Q_i]^{\mathrm{T}} \tag{1}$$

where $U_i, P_i, and\ Q_i$ denote the voltage magnitude, active power injection, and reactive power injection, respectively, and $(\cdot)^T$ denotes the transpose operation. The overall feature matrix is $\boldsymbol{X} \in \mathbb{R}^{N\times3}$.

The fault location task is formulated through a learnable mapping function:

$$f\colon (\boldsymbol{X}, \boldsymbol{G}, \boldsymbol{Z}) \to y \tag{2}$$

where $y \in \{1, \dots, N\}$ denotes the index of the faulted node.

To address the limitations of existing approaches, a Dual-Graph Architecture Physics-Informed Graph Attention Network (DG-PIGAT) is proposed. The model decouples physical connectivity and electrical coupling by constructing a dual-graph structure, and incorporates KCL-related residual information and weak physical regularization to improve physical consistency during training.

Figure 1 presents the overall DG-PIGAT architecture. The framework can be organized into four main stages: input and preprocessing, dual-graph construction, dual-branch feature extraction, and feature fusion with physical constraints.

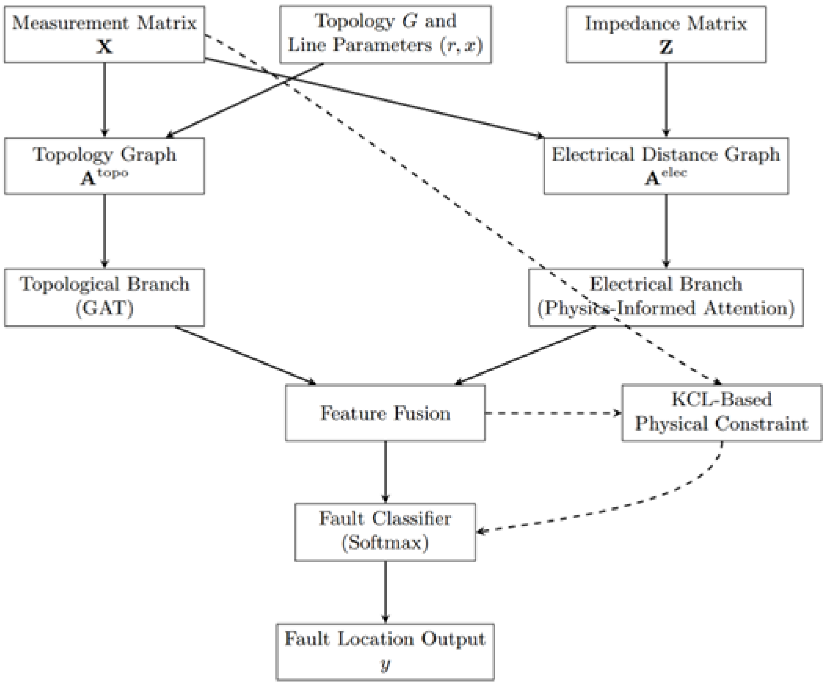

Figure 1. Overall framework of the proposed Dual-Graph Architecture Physics-Informed Graph Attention Network (DG-PIGAT)

The framework consists of four main stages: input preprocessing, dual-graph construction, dual-branch feature extraction, and feature fusion with physical constraint. Specifically, measurement features, network topology, and the impedance matrix are first obtained as inputs. The model then builds a dual-graph representation, including a static topology graph and an electrical distance graph. Based on the two graphs, dual-branch graph attention networks are employed to extract high-order fault features, where the electrical branch uses electrical-distance-guided attention. Finally, the extracted features are fused with KCL-related residual information, and a weak KCL-based regularization term is used to improve the physical consistency of the fault-location results.

3.2. Dual-Graph Construction

3.2.1. Static Topology Graph

The static topology graph is constructed from the feeder connectivity of the distribution network. It provides the physical neighborhood used by the topology branch and serves as a reference for enforcing KCL-related physical consistency.

The static topology graph captures the physical connectivity of the distribution network. Its adjacency matrix $\boldsymbol{A}^{topo} \in \{0,1\}^{N\times N}$ is defined as:

$$A_{ij}^{topo} = \begin{cases} 1, & if\ nodes\ i\ and\ j\ are\ directly\ connected \\ 0, & otherwise \end{cases} . \quad (3)$$

The topology graph is independent of operating conditions and provides fundamental physical connectivity constraints for the model.

3.2.2. Dynamic Electrical Distance Graph

Because the static topology graph cannot quantify electrical coupling strength, a dynamic electrical distance graph $G_{elec} = (V, E_{elec})$ is built from the node impedance matrix.

First, the node impedance matrix $\boldsymbol{Z} \in \mathbb{R}^{N\times N}$ is obtained from network topology and line parameters. The electrical distance (ED) for node pair i and j is given by:

$$ED_{ij} = |Z_{ii} - Z_{ij}| \quad (4)$$

where Z_{ii} denotes the self-impedance of node i, and Z_{ij} denote the mutual impedances associated with nodes i and j.

The electrical distance reflects the sensitivity of node i's voltage to current injection at node j. A smaller value of ED_{ij} indicates stronger electrical coupling and higher correlation of fault disturbances.

To quantify electrical coupling strength, a Gaussian kernel is applied to map electrical distance into continuous adjacency weights:

$$A_{ij}^{init} = exp\left(-\frac{ED_{ij}^2}{2\sigma^2}\right) \quad (5)$$

where σ is the bandwidth parameter controlling the decay rate.

To reduce computational complexity and eliminate weak couplings, a threshold d_{th} is applied:

$$A_{ij}^{elec} = \begin{cases} A_{ij}^{init}, & A_{ij}^{init} \geq d_{th} \\ 0, & othewise \end{cases}. \tag{6}$$

The resulting adjacency matrix defines the edge set E_{elec}, which is used for message passing in the electrical branch.

The electrical distance graph is adaptive to system changes. When topology reconfiguration or parameter variations occur, the impedance matrix is updated accordingly, enabling the model to capture nonlinear coupling variations. Even without topology changes, fault-induced impedance variations can also be reflected through updated electrical distances.

3.3. Dual-Branch Feature Extraction

To capture complementary information from network structure and electrical coupling, two parallel graph attention network (GAT) branches are constructed. The first branch operates on the topology graph, while the second branch is defined on the electrical distance graph.

3.3.1. Topological Branch

The topological branch aggregates node features within the topological neighborhood $\mathcal{N}_i^{topo}$, and its output is expressed as:

$$\boldsymbol{h}_i^{topo} = \sigma\left(\sum_{j\in\mathcal{N}_i^{topo}} \alpha_{ij}^{topo} \boldsymbol{W}^{topo} \boldsymbol{x}_j\right), \tag{7}$$

where $\boldsymbol{x}_j$ is the input feature of node j, $\mathcal{N}_i^{topo}$ represents the topological neighborhood of node i, $\boldsymbol{W}^{topo}$ is a learnable linear transformation matrix, α_{ij}^{topo} is the normalized attention coefficient, and $\sigma(\cdot)$ is a nonlinear activation function.

3.3.2. Electrical Branch

Unlike the topological branch, the electrical branch explicitly incorporates electrical coupling into the attention calculation. The corresponding attention coefficient is defined as:

$$\alpha_{ij}^{elec} = \frac{exp\left(\text{LeakyReLU}(\mathbf{a}^{elec\top}[\boldsymbol{W}^{elec}\boldsymbol{x}_i \| \boldsymbol{W}^{elec}\boldsymbol{x}_j]) \cdot A_{ij}^{elec}\right)}{\sum_{k\in\mathcal{N}_i^{elec}}^{\Sigma(\text{LeakyReLU}(\mathbf{a}^{elec\top}[\boldsymbol{W}^{elec}\boldsymbol{x}_i \| \boldsymbol{W}^{elec}\boldsymbol{x}_k])\cdot A_{ik}^{elec})} exp} \tag{8}$$

Here a^{elec} denotes a learnable attention vector, and $\|$ indicates the concatenation operation, A_{ij}^{elec} represents the electrical adjacency weight, and $\mathcal{N}_i^{elec}$ denotes the electrical neighborhood of node i.

By introducing the electrical adjacency weight into the attention calculation, nodes with stronger electrical coupling are assigned greater importance during feature aggregation. This design enables the model to better capture the propagation characteristics of fault disturbances, especially when the topological distance does not accurately reflect the actual electrical influence between nodes. The use of electrical-coupling-aware weighting is also consistent with previous studies on power-grid structural analysis, where electrical parameters have been incorporated to characterize the functional relationships among buses rather than relying solely on topological connectivity [28].

3.4. Feature Fusion and Physical Constraint

3.4.1. Feature Fusion and Classification

The two branch outputs are concatenated and projected:

$$\boldsymbol{H}^{final} = \text{ReLU}\left(\boldsymbol{W}_f[\boldsymbol{H}^{topo} \| \boldsymbol{H}^{elec}] + \boldsymbol{b}_f\right), \tag{9}$$

where $\boldsymbol{H}^{topo}$ and $\boldsymbol{H}^{elec}$ denote the feature representations from the two branches, $\boldsymbol{W}_f$ and $\boldsymbol{b}_f$, are learnable parameters.

The fused feature is used to produce a global fault-location score over all buses. In addition, KCL-related and electrical residual information can be used to enhance node-level fault scoring, so that the final prediction reflects both learned graph features and physically meaningful local anomalies.

3.4.2. KCL-Based Physical Consistency Constraint

To alleviate the black-box nature of data-driven models, Kirchhoff's Current Law (KCL) is incorporated into the learning process through residual feature construction and a physical regularization term. For any node, the current balance condition is expressed as follows:

$$\mathcal{L}_{phy} = \frac{1}{N}\sum_{i=1}^{N}\left|\sum_{j\in\mathcal{N}_i^{topo}} \frac{U_i - U_j}{z_{ij}} - \frac{\sqrt{P_i^2 + Q_i^2}}{U_i}\right|, \tag{10}$$

where $\mathcal{N}_i^{topo}$ denotes the directly connected neighbor set of node i, and represents the line impedance linking nodes i and j.

The first term represents the total current flowing into or out of node i through all connected lines, while the second term denotes the magnitude of the load current at node j. A smaller absolute difference between the two terms indicates better compliance with the KCL current conservation law, and thus leads to a smaller physical loss.

Meanwhile, the task loss for fault location is defined by the cross-entropy function:

$$\mathcal{L}_{task} = -\frac{1}{N}\sum_{i=1}^{N} y_i\, log(\hat{y}_i), \tag{11}$$

where y_i gives the true label of the faulted node, and $\hat{y}_i$ denotes the model-estimated probability output by the model.

The final training objective is written as a weighted combination of the task loss and the physical regularization loss:

$$\mathcal{L}_{total} = \mathcal{L}_{task} + \lambda \mathcal{L}_{phy}, \tag{12}$$

where λ is the weighting coefficient used to balance task fitting accuracy and physical consistency.

3.5. Complexity Analysis

For the topological branch, the time complexity is $O(M \cdot F)$, where M denotes the number of edges and F is the feature dimension. The electrical branch has complexity $O(N \cdot k \cdot F)$, where k is the average number of electrical neighbors. Since distribution networks typically satisfy $M \approx N$, the overall time complexity can be approximated as $O(N \cdot F)$.

In terms of memory consumption, the dominant cost arises from the impedance matrix and the learnable model parameters. The resulting space complexity is therefore $O(N^2 + F^2)$.

4. Experiments

4.1. Experimental Setup

4.1.1. Test System

The IEEE 33-bus radial distribution system is adopted as the test platform. This benchmark feeder was originally introduced for distribution network reconfiguration studies and has since been widely used in distribution system analysis [29]. It has a base voltage of 12.66 kV, a total load of 3.715 MW and 2.3 MVar, 33 buses, and 32 distribution lines.

4.1.2. Dataset Generation

A simulation model is established in MATLAB/Simulink to generate fault samples in batches. Four typical scenarios are considered, each corresponding to a different evaluation purpose. S1 evaluates baseline localization capability under mild disturbances; S2 examines weak fault-signature recognition under high-impedance faults; S3 evaluates generalization under fixed DER integration; and S4 is designed to test the core topology-electrical mismatch condition where the electrical graph is expected to provide complementary information beyond topology-based message passing.

S1: Normal condition with slight perturbations. No DER integration and fixed topology are considered. The fault resistance ranges from 0.5 to 8 Ω. Voltage magnitude noise is set to 6%, phase angle noise to 3%, and 2% feature missing is introduced to simulate slight disturbances in normal operation.

S2: High-impedance fault scenario. No DER integration and fixed topology are considered. The fault resistance ranges from 15 to 40 Ω. Voltage magnitude noise is 12%, phase angle noise is 6%, and 6% feature missing is introduced to simulate high-impedance grounding faults.

S3: Fixed-DER integration scenario. Distributed generation is connected at nodes 13, 17, 22, and 28, while the feeder topology remains fixed. The fault resistance ranges from 7 to 28 Ω. Voltage magnitude noise is 8%, phase angle noise is 4%, and 3% feature missing is introduced. In addition, fixed DER voltage support and remote electrical-coupling responses are included to simulate active distribution network operation in which the physical topology remains unchanged, but fault signatures may also propagate through electrically coupled non-neighbor buses.

S4: Topology-electrical mismatch scenario. DER integration with varying connection locations and topology-related measurement disturbance is considered. The fault resistance ranges from 6 to 22 Ω. Voltage magnitude noise is 7%, phase angle noise is 3.5%, and 4% feature missing is introduced. In addition, measurements associated with selected topological neighbors are disturbed, and local voltage support may partially mask the voltage drop at the faulted node, while electrically coupled non-neighbor nodes retain coherent fault-related responses. This setting simulates cases where physical topology does not fully reflect electrical coupling under varying active distribution network conditions.

Three fault categories are considered: 0 (single-phase-to-ground fault), 1 (line-to-line fault), and 2 (three-phase fault). A total of 2400 samples are generated for each scenario, resulting in 9600 samples in total. The data are split into training and testing subsets using an 8:2 ratio.

4.1.3. Dataset Generation

For comparison, the following representative baseline methods are used:

1. Traditional GNN: A graph convolutional network based on static topology, serving as a classical baseline in distribution network perception tasks [8];
2. Single Graph GAT: A standard graph attention network constructed on static topology, representing mainstream GNN-based approaches for fault location [10];
3. GraphSAGE: An inductive graph neural network baseline based on neighborhood aggregation, used to represent strong spatial aggregation methods [11];
4. GATv2: A dynamic graph attention baseline that addresses the static-attention limitation of the original GAT and provides a stronger attention-based comparator [12];
5. ChebNet: A spectral graph convolution baseline based on Chebyshev polynomial approximation [13].
6. DG-PIGAT (Ours): The proposed dual-graph physics-informed graph attention network.

4.1.4. Evaluation Metrics

The main metric is fault-location accuracy (Accuracy), computed as:

$$\text{Accuracy} = \frac{\textit{Number of correctly located fault samples}}{\textit{Total number of test samples}} \times 100\%. \tag{13}$$

In addition, repeated experiments with five random seeds are conducted. Mean accuracy, standard deviation, F1-score, inference time, and paired t-test p-values against DG-PIGAT are reported. Convergence curves, confusion matrix, sensitivity curves, robustness profiles, and ablation studies are also used to evaluate optimization behavior, per-bus error distribution, sensitivity to measurement quality, and module contribution.

4.1.5. Model Configuration

The principal model settings are configured as follows: the proposed graph attention backbone uses a hidden dimension of 128, the electrical graph is constructed from impedance-based electrical distance, and optimization is performed with Adam at an initial learning rate of 8×10^{-4}, batch size of 256, and 200 training epochs. Unless otherwise specified, repeated experiments use five random seeds (42-46). During training, the physics regularization weight is gradually increased to 0.001, while the voltage-anchor weight is set to 0.001. All experiments are conducted using Python 3.10, PyTorch, and PyTorch Geometric.

4.2. Basic Experimental Results and Convergence Analysis

4.2.1. Performance under Normal Conditions with Slight Perturbations

Table 1 lists the repeated fault-location results under Scenario S1. As S1 is a mild operating condition, most advanced GNN methods yield satisfactory performance.

Table 1 Fault location accuracy under normal conditions with slight perturbations

Method	Accuracy (%)
Traditional GNN	40.88 ± 4.08
Single Graph GAT	98.75 ± 0.33
GraphSAGE	99.17 ± 0.42
GATv2	99.04 ± 0.48
ChebNet	98.25 ± 0.98
DG-PIGAT (Ours)	99.50 ± 0.24

From Table 1, DG-PIGAT obtains 99.50% accuracy under normal operation, which is comparable to Single Graph GAT, GraphSAGE, and GATv2. The small differences among these high-performing methods are not statistically significant, indicating that the proposed dual-graph and physics-informed design does not sacrifice baseline performance in normal operating conditions.

4.2.2. Convergence Analysis

Figure 2 shows the loss trajectory and test accuracy of Single Graph GAT, GraphSAGE, and DG-PIGAT under the topology-electrical mismatch scenario (S4).

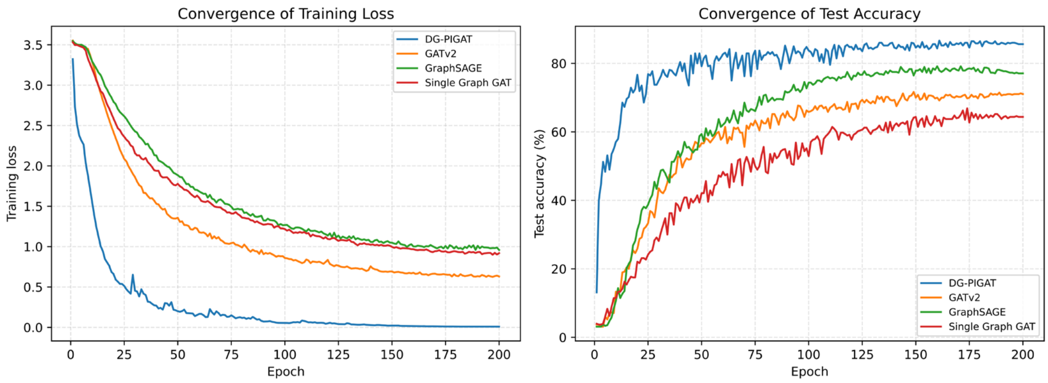

Figure 2. Convergence curves of different methods under the topology-electrical mismatch scenario

The convergence curves are used to evaluate optimization stability rather than only final accuracy. Under S4, DG-PIGAT achieves the highest diagnostic accuracy among the compared methods (85.63% for the representative seed), while GraphSAGE, GATv2, and Single Graph GAT obtain 77.08%, 71.04%, and 64.38%, respectively. This result shows that the proposed model can maintain effective optimization under topology-electrical mismatch, where topology-only message passing can be misleading.

4.3. Performance under Stress Scenarios

Table 2 summarizes the repeated fault-location accuracy of six methods under the four consolidated scenarios. Each value is reported as mean +/- standard deviation over five random seeds, and paired t-tests are conducted against DG-PIGAT.

Table 2. Fault location accuracy under different scenarios

Scenario	Traditional GNN	Single Graph GAT	GraphSAGE	GATv2	ChebNet	DG-PIGAT
S1	40.88 ± 4.08	98.75 ± 0.33	99.17 ± 0.42	99.04 ± 0.48	98.25 ± 0.98	99.50 ± 0.24
S2	6.00 ± 2.28	76.25 ± 3.18	74.79 ± 3.05	82.50 ± 1.69	58.96 ± 3.04	86.17 ± 1.63
S3	10.33 ± 1.94	83.88 ± 5.06	88.54 ± 2.61	89.29 ± 0.64	80.21 ± 2.26	92.67 ± 1.90
S4	17.96 ± 6.81	66.08 ± 2.63	78.63 ± 3.39	75.50 ± 3.44	65.67 ± 2.99	86.92 ± 1.64

The results show that DG-PIGAT is competitive under S1¬–S3 and achieves the best performance under S4, which is the scenario most aligned with the motivation of topology-electrical inconsistency. Under S4, DG-PIGAT reaches 86.92% accuracy, outperforming GraphSAGE by 8.29 percentage points (p=0.0025) and GATv2 by 11.42 percentage points (p=0.0025). GraphSAGE performs strongly under S1 and S3, indicating that neighborhood aggregation remains effective in regular or fixed-DER conditions. Therefore, the main advantage of DG-PIGAT should be interpreted as improved robustness under topology-electrical mismatch and degraded measurement conditions, rather than universal dominance in every scenario.

The bus-level error distribution under this scenario is further examined through the normalized confusion matrix in Figure 3.

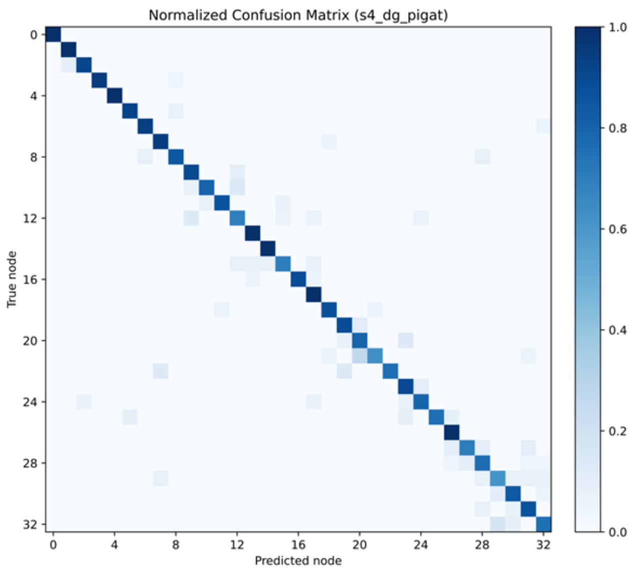

Figure 3. Normalized confusion matrix of DG-PIGAT under the topology-electrical mismatch scenario

The confusion matrix is used to examine the distribution of fault-location errors over the 33 buses. A strong diagonal indicates that most fault buses are correctly localized, while off-diagonal entries reveal which buses are easily confused.

Under S4, DG-PIGAT shows no systematic failure over a particular feeder section. Misclassifications mainly occur among electrically similar or adjacent buses, especially near long feeder ends. This pattern indicates stable per-bus localization under topological electrical inconsistency.

4.4. Ablation Study

The ablation study focuses on the topology-electrical mismatch scenario S4, because this scenario is specifically designed to evaluate whether the electrical-distance branch provides complementary information when topology-based neighborhoods become partially misleading. The results are reported in Table 3.

Table 3. Ablation study under the topology-electrical mismatch scenario

Variant	Accuracy (%)
Single Graph GAT	66.33 ± 2.72
DG topology branch only	84.88 ± 1.18
DG w/o KCL physics	86.38 ± 2.04
DG-PIGAT	86.71 ± 1.34

The ablation results show that the DG variants substantially outperform the single-graph GAT baseline, confirming that the dual-graph architecture provides a stronger representation under topology-electrical mismatch. Compared with the topology-branch-only variant, DG-PIGAT improves the accuracy from 84.88% to 86.71%, indicating that electrical-distance-guided propagation contributes additional fault-location information when topological neighborhoods are partially misleading. The model variant excluding KCL-related physics yields a marginally higher classification accuracy (86.38 ± 2.04 vs. 86.71 ± 1.34). This subtle performance difference indicates that the KCL component cannot be regarded as a universal accuracy-boosting module. Instead, it introduces physical inductive bias and supports physically plausible learning while keeping the full model competitive.

4.5. Sensitivity Analysis

To evaluate the influence of measurement quality, noise amplitude, and missing-feature rate are independently varied under the topology-electrical mismatch test condition. Figure 4 **s**hows the corresponding sensitivity curves, and representative numerical results are summarized in Table 4.

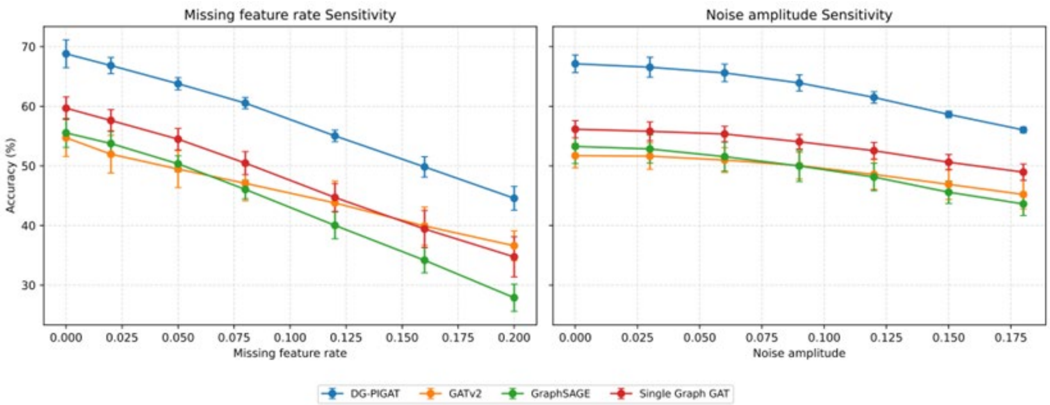

Figure 4. Measurement-noise and missing-feature sensitivity curves under topology-electrical mismatch

Table 4. Representative sensitivity results under topology-electrical mismatch

Type	Level	Single Graph GAT	GraphSAGE	GATv2	DG-PIGAT	Gain over the best baseline
noise_amp	0.00	56.15	53.27	51.72	67.13	10.98
noise_amp	0.09	54.05	49.98	50.05	63.92	9.87
noise_amp	0.18	48.95	43.62	45.20	56.03	7.08
missing_rate	0.00	59.68	55.53	54.72	68.80	9.12
missing_rate	0.12	44.70	40.02	43.78	55.03	10.33
missing_rate	0.20	34.73	27.88	36.60	44.57	7.97

DG-PIGAT consistently achieves the highest accuracy among the compared methods across the tested noise and missing-data levels. In Table 4, Gain over best baseline denotes the accuracy improvement of DG-PIGAT relative to the strongest non-DG-PIGAT result at each degradation level. Across the representative degradation levels, DG-PIGAT improves accuracy over the strongest baseline by 7.08 to 10.98 percentage points, indicating better tolerance to measurement degradation.

4.6. Robustness Test

Robustness is further evaluated using multiple stress profiles, including high-impedance faults, fixed DER integration, topology-electrical mismatch, heavy noise, heavy missing data, and combined noise-missing disturbance. The results are summarized in Table 5.

Table 5. Robustness results under different stress profiles

Profile	Single Graph GAT	GraphSAGE	GATv2	DG-PIGAT	Gain over the best baseline
normal	98.00	98.33	99.03	98.73	- 0.30
high_impedance	64.35	68.08	65.53	71.38	3.30
der_fixed_nodes	83.08	87.28	83.6	88.17	0.88
topology_mismatch	55.00	51.92	50.25	65.90	10.90
heavy_noise	56.53	61.38	58.87	64.07	2.68
heavy_missing	20.13	27.78	36.95	37.70	0.75
noise_and_missing	31.60	38.55	44.40	45.90	1.50
weak_measurement	71.75	76.63	73.82	77.43	0.80

The robustness results are mixed but informative. In Table 5, Gain over best baseline denotes the accuracy improvement of DG-PIGAT relative to the strongest non-DG-PIGAT method under each stress profile. DG-PIGAT achieves the best result under high-impedance faults, fixed-DER operation, topology-electrical mismatch, heavy noise, heavy missing data, combined noise-missing disturbance, and weak measurement conditions. GATv2 is slightly higher only under the normal profile, where all advanced GNN methods achieve similarly high accuracy. These results suggest that DG-PIGAT is especially beneficial when measurement quality degrades or topology-based neighborhoods become less reliable.

4.7. Practical Distribution-Network Case Based on Iowa 240-Bus Real Data

The Iowa 240-bus Distribution Test System [30] serves as a high-fidelity environment for evaluating DG-PIGAT's resilience against the complexities of utility-scale operations. Unlike simplified benchmarks, this model comprises three expansive feeders and utilizes a full year of actual smart-meter records to capture authentic load volatility.

Because the public Iowa dataset does not provide labeled field fault records, this study does not claim a direct field fault-location experiment. Instead, the real feeder topology, line parameters, and 8760 hourly P/Q load profiles are used to construct time-varying operating conditions, and labeled fault samples are generated by fault injection under these real load profiles. The first 75% of the hourly records are used for training-condition sampling, whereas the remaining 25% are used for testing-condition sampling, so that the test set reflects unseen seasonal operating periods. Measurement noise, partial feature missing, local voltage support, and topology-electrical coupling effects are included to approximate degraded practical measurements.

Six baseline methods are compared under the same data split and training setting. For the Iowa case, exact Top-1 bus identification is difficult because the task becomes a 240-class localization problem, and adjacent buses may show similar voltage-disturbance patterns. Therefore, both exact accuracy and topology-neighborhood localization accuracy are reported. Neighbor-1 and Neighbor-3 accuracy denote whether the predicted bus falls within one or three topological hops of the actual fault bus, respectively.

Table 6. Real-data-driven validation on the Iowa 240-bus distribution system

Method	Top-1 accuracy (%)	Neighbor-1 accuracy (%)	Neighbor-3 accuracy (%)
Traditional GNN	10.20 ± 0.64	26.53	51.87
Single Graph GAT	8.27 ± 2.29	21.93	46.83
GraphSAGE	8.87 ± 2.31	23.30	48.47
GATv2	7.13 ± 0.76	19.50	42.13
ChebNet	8.77 ± 1.00	25.03	50.47
DG-PIGAT (Ours)	17.93 ± 1.64	40.10	66.80

The comparative results in Table 6 reveal a substantial performance gap between DG-PIGAT and conventional GNN architectures. While the expanded search space naturally depresses the Top-1 accuracy across all models, DG-PIGAT maintains a 17.93% exact match rate, which is substantially higher than the 0.42% random-guess level for a 240-class task. More importantly, the model achieves a Neighbor-3 accuracy of 66.80%, outperforming the strongest baseline by 14.93 percentage points. This performance gain suggests that when fault signatures diffuse throughout a large-scale feeder, the electrical-coupling branch in our architecture acts as a "physical anchor," successfully localizing the fault within a manageable neighborhood even when pure topological message-passing becomes ambiguous.

While these findings stem from a high-fidelity simulation rather than live field deployment, the use of genuine feeder geometries and smart-meter profiles provides a much-needed stress test for graph-based models. This case study effectively complements the IEEE 33-bus experiments by confirming that DG-PIGAT's advantage is not merely a byproduct of small-scale network properties, but a scalable trait that holds under the structural and temporal complexities of real-world utility networks.

4.8. Experimental Conclusions

The experimental results lead to the following observations:

1. Under the normal operating scenario, DG-PIGAT reaches a performance level comparable to strong GNN baselines, indicating that the dual-graph and physics-informed components do not noticeably reduce baseline localization accuracy.
2. The proposed method shows its most evident benefit in the topology-electrical mismatch scenario. In S4, DG-PIGAT achieves 86.92% accuracy and significantly outperforms GraphSAGE and GATv2, suggesting that electrical-coupling information is useful when topology-only neighborhoods become partially misleading.
3. The sensitivity and robustness results show that DG-PIGAT is less affected by measurement noise and missing features than GraphSAGE and Single Graph GAT. This indicates that the proposed design is better suited to degraded measurement conditions.
4. The ablation study under S4 indicates that the electrical branch contributes additional information under topology-electrical mismatch. The KCL-related component is interpreted conservatively: it introduces physical inductive bias and supports physically plausible learning, but it does not uniformly maximize classification accuracy.

5. Conclusion

This paper proposes DG-PIGAT, a dual-graph physics-informed graph attention network for fault localization in active distribution networks. The method jointly models feeder topology and impedance-related electrical coupling, and introduces KCL-related residual information to guide physically consistent learning. In this way, DG-PIGAT aims to address cases where topology-based message passing alone may be insufficient.

Experiments on the IEEE 33-bus distribution system with four consolidated scenarios and six comparison methods show that DG-PIGAT remains competitive under normal, high-impedance, and DER-integration conditions. The most distinctive improvement is observed under the topology-electrical mismatch. In this scenario, DG-PIGAT reaches 86.92% accuracy and significantly outperforms GraphSAGE by 8.29 percentage points ($p=0.0025$) and GATv2 by 11.42 percentage points ($p=0.0025$). Additional sensitivity and robustness analyses show that the method has better tolerance to measurement noise, missing features, and combined disturbances. These results support the value of combining topology information, electrical coupling, and weak physical regularization for robust fault localization.

While these results support the value of dual-graph modeling, the transition from simulation studies to field deployment remains an important next step. Future research will focus on scaling the architecture to larger and more heterogeneous distribution systems and validating its performance using real-world field data. In addition, addressing extreme data sparsity and developing adaptive, lightweight versions of the model will be important for its integration into real-time edge-computing devices in dynamic grid environments.

Acknowledgements

The authors would like to thank the distribution network operation and maintenance engineers and technical specialists for their helpful discussions on practical fault localization requirements, operating scenarios, and data platform constraints.

Author Contributions

Conceptualization, Y.Z. and Y.L.; methodology, Y.Z. and P.Q.; software, R.Z.; validation, Y.Z., R.Z. and Z.S.; formal analysis, P.Q.; investigation, R.Z. and Y.Zha.; resources, Z.S.; data curation, Y.Zha.; writing—original draft preparation, Y.Z.; writing—review and editing, Y.L.; supervision, Y.L.; project administration, Y.L. All authors have read and agreed to the published version of the manuscript.

Funding

This research was supported by the Science and Technology Project of State Grid Sichuan Electric Power Company, "Research on Key Technologies for High-Speed Heterogeneous Data Aggregation and Joint Intelligent Labeling for Distribution Network Operation" (No. 521947250004).

Ethics Statement

Not applicable

Data Availability Statement

The data used in this study were generated through simulation and are not publicly available.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] Arefifar, S.A.; Alam, M.S.; Hamadi, A. A review on self-healing in modern power distribution systems. Journal of Modern Power Systems and Clean Energy 2023, 11, 1719–1733.

[2] Zidan, A.; Khairalla, M.; Abdrabou, A.M.; et al. Fault detection, isolation, and service restoration in distribution systems: State-of-the-art and future trends. IEEE Transactions on Smart Grid 2016, 8, 2170–2185.

[3] Mora-Florez, J.; Meléndez, J.; Carrillo-Caicedo, G. Comparison of impedance-based fault location methods for

power distribution systems. Electric Power Systems Research 2008, 78, 657–666.
[4] Xie, L.; Luo, L.; Li, Y.; et al. A traveling wave-based fault location method employing VMD-TEO for distribution network. IEEE Transactions on Power Delivery 2019, 35, 1987–1998.
[5] Liu, J.; Qu, Q.; Yang, H.; et al. Deep learning-based intelligent fault diagnosis for power distribution networks. International Journal of Computers Communications & Control 2024, 19.
[6] Xiao, M.; Li, Y.; Zhou, Z.; et al. Distribution network fault location model based on DCNN-BiLSTM. Journal of Physics: Conference Series 2025, 3087, 012026.
[7] Li, K.; Yu, X.; Ou, S.; et al. Fault location in distribution network based on RNN and transfer learning. International Journal of High Speed Electronics and Systems 2025, 34.
[8] Kipf, T.N.; Welling, M. Semi-supervised classification with graph convolutional networks. In Proceedings of the 5th International Conference on Learning Representations, Toulon, France, 24–26 April 2017.
[9] Chen, K.; Hu, J.; Zhang, Y.; Yu, Z.; He, J. Fault location in power distribution systems via deep graph convolutional networks. IEEE J. Sel. Areas Commun. 2020, 38, 119–131.
[10] Veličković, P.; Cucurull, G.; Casanova, A.; Romero, A.; Liò, P.; Bengio, Y. Graph attention networks. In Proceedings of the 6th International Conference on Learning Representations, Vancouver, BC, Canada, 30 April–3 May 2018.
[11] Hamilton, W.L.; Ying, R.; Leskovec, J. Inductive representation learning on large graphs. Advances in neural information processing systems, 2017, 30.
[12] Brody, S.; Alon, U.; Yahav, E. How attentive are graph attention networks? International Conference on Learning Representations, 2022.
[13] Defferrard, M.; Bresson, X.; Vandergheynst, P. Convolutional neural networks on graphs with fast localized spectral filtering. Advances in neural information processing systems, 2016, 29.
[14] Liu, J.; Huang, Y.; Chen, K.; et al. Graph neural networks for fault diagnosis in photovoltaic-integrated distribution networks with weak features. Sensors 2025, 25, 5691.
[15] Karabulut, B.; Manna, C.; Develder, C. Generalization of graph neural network models for distribution grid fault detection. In Proceedings of the 2025 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm); IEEE, 2025; pp. 1–7.
[16] Wang, Z.; Huang, B.; Zhou, B.; et al. An enhanced fault localization technique for distribution networks utilizing cost-sensitive graph neural networks. Processes 2024, 12, 2312.
[17] Liang, L.; Zhang, H.; Cao, S.; et al. Fault location method for distribution networks based on multi-head graph attention networks. Frontiers in Energy Research 2024, 12, 1395737.
[18] Song, X.; Chen, C.; Yan, X.; et al. KG-FLoc: Knowledge graph-enhanced fault localization in secondary circuits via relation-aware graph neural networks. Electronics 2025, 14.
[19] Donon, B.; Donnot, B.; Guyon, I.; et al. Graph neural solver for power systems. In Proceedings of the 2019 International Joint Conference on Neural Networks (IJCNN); IEEE, 2019; pp. 1–8.
[20] Raissi, M.; Perdikaris, P.; Karniadakis, G.E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational Physics 2019, 378, 686–707.
[21] Karniadakis, G.E.; Kevrekidis, I.G.; Lu, L.; Perdikaris, P.; Wang, S.; Yang, L. Physics-informed machine learning. Nature Reviews Physics 2021, 3, 422–440.
[22] Zamzam, A.S.; Sidiropoulos, N.D. Physics-aware neural networks for distribution system state estimation. IEEE Transactions on Power Systems 2020, 35, 4347–4356.
[23] Misyris, G.S.; Venzke, A.; Chatzivasileiadis, S. Physics-informed neural networks for power systems. In Proceedings of the 2020 IEEE Power & Energy Society General Meeting (PESGM); IEEE, 2020; pp. 1–5.
[24] Wu, H.; Xu, Z.; Wang, M.; et al. Meta-learning enhanced physics-informed graph attention convolutional network for distribution power system state estimation. IEEE Transactions on Network Science and Engineering 2025, 12, 1186–1197.
[25] Li, W.; Deka, D. PPGN: Physics-preserved graph networks for real-time fault location in distribution systems with limited observation and labels. In Proceedings of the Annual Hawaii International Conference on System Sciences; 2023.
[26] Jin, H.; Balaprakash, P.; Zou, A.; et al. Physics-informed heterogeneous graph neural networks for DC blocker placement. Electric Power Systems Research 2024, 235, 110795.
[27] Okoyomon, E.; Yaniv, A.; Goebel, C. Physics-informed inductive biases for voltage prediction in distribution grids. arXiv 2025, arXiv:2509.25158.
[28] Wang, Z.; Scaglione, A.; Thomas, R.J. Electrical centrality measures for electric power grid vulnerability analysis. In Proceedings of the 49th IEEE Conference on Decision and Control, Atlanta, GA, USA, 15–17 December 2010; pp. 5792–5797.
[29] Baran, M.E.; Wu, F.F. Network reconfiguration in distribution systems for loss reduction and load balancing. IEEE Trans. Power Deliv. 1989, 4, 1401–1407.
[30] Bu, F.; Yuan, Y.; Wang, Z.; Dehghanpour, K.; Kimber, A. A Time-series Distribution Test System based on Real Utility Data. In Proceedings of the 2019 North American Power Symposium (NAPS), Wichita, KS, USA, 2019; pp. 1–6.