

Article

A Dual-View Mixup-ResNet Method for Intelligent Monitoring and Fault Diagnosis of Cable Sheath Circulating Current Signals

Haiqi Yang ^{1,4} , Jinwei Mao ², Bo Zhang ^{3,4}, Jize He ^{1,4} and Xiaoyu Liang ^{1,*} 

¹ College of Energy, Environment and Safety Engineering, China Jiliang University, Hangzhou 310018, China; s25060837018@cjl.u.edu.cn (H.Y.)

² Shaoxing Shangyu Hangxie Thermal Power Co., Ltd., Shaoxing 312000, China

³ College of Metrology Measurement and Instrument, China Jiliang University, Hangzhou 310018, China

⁴ Joint Innovation Key Laboratory of Ministry of Emergency Management for Emerging Risk Identification, Prevention and Control in Safety Production, Beijing 100020, China

* Correspondence: xyliang@cjl.u.edu.cn; Tel.: +86-13777846029

Abstract

In renewable-powered distribution systems and microgrids, reliable cable condition monitoring is essential for operational security and early fault detection. Sheath circulating current signals provide valuable information for identifying grounding abnormalities and incipient faults, but their diagnosis is difficult because the signals exhibit strong inter-phase coupling, and fault samples are limited in practice. This study proposes a dual-view Mixup-ResNet framework for fault diagnosis of cable sheath circulating current signals. Specifically, physical-range normalization is employed to retain magnitude-related fault information and inter-phase proportional relationships, while sample-wise z-score normalization is used to emphasize waveform morphology. These two complementary views are concatenated to form a six-channel input for a lightweight one-dimensional residual network, which is trained with Mixup, label smoothing, dropout, and cosine annealing. On an ATP-EMTP-generated eight-class dataset, the proposed method achieves an average accuracy of 91.50%, a weighted F1-score of 91.69%, and a macro F1-score of 91.69% under a unified 5×5 repeated stratified cross-validation protocol. Additional tests under 12% relative Gaussian noise show that the method maintains competitive Gaussian-noise tolerance within the tested condition, although broader field disturbances remain to be further validated. These findings suggest that the proposed method has potential for small-sample fault diagnosis of cable sheath circulating current signals and provides a preliminary basis for intelligent cable condition monitoring.

Keywords: power cable; sheath circulating current; condition monitoring; fault diagnosis; small-sample diagnosis; Mixup; one-dimensional residual network; renewable-powered microgrids



Academic Editor: Dragos Pasculescu

Received: 19 May 2026

Revised: 25 June 2026

Accepted: 26 June 2026

Published: 13 July 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and

conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

With the increasing penetration of renewable energy and the rapid development of distributed generation, energy storage, and flexible loads, modern power systems are operating under stronger fluctuations and tighter coupling [1]. Under such circumstances, condition monitoring, anomaly identification, and early fault warning of critical electrical assets have become increasingly important for maintaining supply reliability and operational resilience [2]. This need is particularly pronounced in renewable-powered

distribution systems and microgrid applications [3]. In these scenarios, conventional monitoring approaches based mainly on manual inspection or simple threshold alarms are often insufficient in terms of accuracy, timeliness, and intelligence [4,5].

High-voltage cross-linked polyethylene (XLPE) cables serve as important transmission carriers in urban distribution networks, industrial power supply systems, and renewable grid integration scenarios [6,7]. The operating condition of their metal sheath grounding system directly affects insulation integrity, current-carrying capability, and long-term service reliability [8,9]. Abnormal sheath circulating currents are commonly associated with insulation degradation, multiple grounding points, local damage, and incipient fault evolution [10,11]. If such anomalies cannot be identified at an early stage, they may further lead to local overheating, accelerated insulation aging, protection maloperation, or even line outage [12,13]. Recent studies have also explored online insulation monitoring based on distributed current analysis for cable-state awareness [14]. Therefore, intelligent monitoring and diagnosis of sheath circulating current signals are valuable for health management, fault warning, and outage risk mitigation [15,16].

Sheath circulating current monitoring is important for both underground urban distribution cable systems and renewable-powered microgrid applications. In underground urban cable systems, abnormal sheath currents are often associated with sheath bonding defects, multiple grounding points, insulation degradation, local overheating, and cable-tunnel operating conditions [17,18]. In renewable-powered microgrids, renewable intermittency, frequent power-electronic switching, inverter-induced harmonics, load fluctuation, and operating-mode changes may introduce more structured disturbances into monitored current signals [19,20]. Therefore, sheath circulating current diagnosis should consider not only conventional cable grounding abnormalities but also the disturbance characteristics introduced by renewable-rich operating environments.

Existing approaches to cable condition monitoring include thermal or optical sensing, multi-point sensor monitoring, and other monitoring-oriented techniques [21,22]. Optical-fiber-assisted and vibration-based partial-discharge diagnostics have also been investigated [23]. Related localization-oriented methods, such as frequency-domain reflectometry, have been reported for cable defect analysis [24]. Compared with optical-fiber-assisted partial-discharge diagnostics and frequency-domain reflectometry, sheath circulating current monitoring provides a more direct indication of the operating state of the metal sheath grounding loop. It is particularly sensitive to abnormal grounding paths, multiple grounding points, sheath bonding defects, and inter-phase coupling effects in the sheath circuit [17,18]. In addition, sheath current signals can be obtained from grounding leads or link-box current sensors, making them suitable as a practical complementary metric for online monitoring rather than a replacement for optical, partial-discharge, or reflectometry-based techniques. In engineering practice, threshold-based alarming and shallow machine learning methods built on handcrafted or mixed-domain features are still widely used [25,26]. Risk-ranking and fuzzy-evaluation schemes have also been reported for cable condition assessment [27]. Although these methods remain useful under certain operating conditions, they still exhibit clear limitations in complex environments [6]. First, sheath circulating current signals are inherently time-dependent, non-stationary, and strongly coupled across phases, making it difficult for conventional empirical features to capture their underlying fault patterns [8,10]. Second, under high renewable penetration, increasing operating variability and disturbance sources impose stricter requirements on robustness, real-time capability, and generalization [1,2]. Third, in practical engineering applications, high-quality fault samples, especially those corresponding to composite faults and incipient abnormal conditions, are typically scarce, which makes data-driven models prone to overfitting under small-sample conditions [16,28].

In recent years, deep learning has provided a promising avenue for intelligent monitoring and fault diagnosis of electrical equipment [2,16]. Similar trends can be observed in recent studies on partial-discharge and cable-defect classification [29,30]. Multi-source fusion schemes have also been explored for real-defect identification [31]. Convolutional neural networks and related architectures can learn discriminative representations directly from raw time-series signals, thereby reducing reliance on complicated handcrafted feature extraction [32,33]. In particular, one-dimensional convolutional models and residual networks are well suited to current and voltage signal analysis [10,34]. However, for sheath circulating current signals, which possess explicit physical magnitudes and inter-phase proportional relationships, using only a statistical normalization strategy may weaken absolute amplitude information related to fault severity [8,9]. Moreover, when the sample size is limited, deep models may still suffer from unstable training and insufficient generalization [16,35]. It is therefore necessary to combine physical priors, robust data augmentation, and lightweight deep architectures in a coordinated manner.

Motivated by these considerations, this paper proposes a dual-view Mixup-ResNet method for intelligent monitoring-oriented fault diagnosis of cable sheath circulating current signals. First, the three-phase sheath current signals are normalized according to the physical measurement range so that fault-related magnitude information and inter-phase proportional relationships can be preserved. Meanwhile, a sample-wise auxiliary z-score view is constructed and concatenated with the physically normalized primary view to form a dual-view input representation, allowing the model to exploit both fault-intensity information and waveform-shape characteristics. On this basis, a lightweight one-dimensional residual network is employed for end-to-end feature extraction and classification, while Mixup data augmentation, label smoothing, dropout, and cosine annealing are jointly introduced to improve classification performance and robustness under small-sample conditions. Experiments on an eight-class dataset generated using ATP-EMTP show that, under a unified 5×5 repeated stratified cross-validation protocol, the proposed method achieves an average diagnostic accuracy of 91.50%, a weighted F1-score of 91.69%, and a macro F1-score of 91.69%. Under 12% relative Gaussian noise, the proposed method obtains an average accuracy of 91.0%, indicating competitive Gaussian-noise tolerance under the tested condition. These results indicate the preliminary feasibility of the proposed method for simulation-based small-sample diagnosis of cable sheath circulating current signals.

The overall workflow of the proposed dual-view Mixup-ResNet framework is illustrated in Figure 1. The main contributions of this paper are summarized as follows.

1. An intelligent monitoring-oriented fault diagnosis framework for cable sheath circulating current signals is developed, providing a preliminary simulation-based approach for condition awareness and early warning studies of critical cable assets in renewable-driven power systems.
2. A dual-view input representation combining a physically normalized primary view and an auxiliary z-score view is constructed to preserve electrical magnitude information while enhancing waveform-shape characterization.
3. A small-sample fault diagnosis method integrating Mixup augmentation with a lightweight 1D-ResNet is designed, and its stability and generalization are further improved through label smoothing, dropout, and cosine annealing.
4. Systematic experiments on a multi-condition ATP-EMTP dataset demonstrate the preliminary feasibility of the proposed method for simulation-based small-sample diagnosis and provide a basis for further laboratory and field validation.

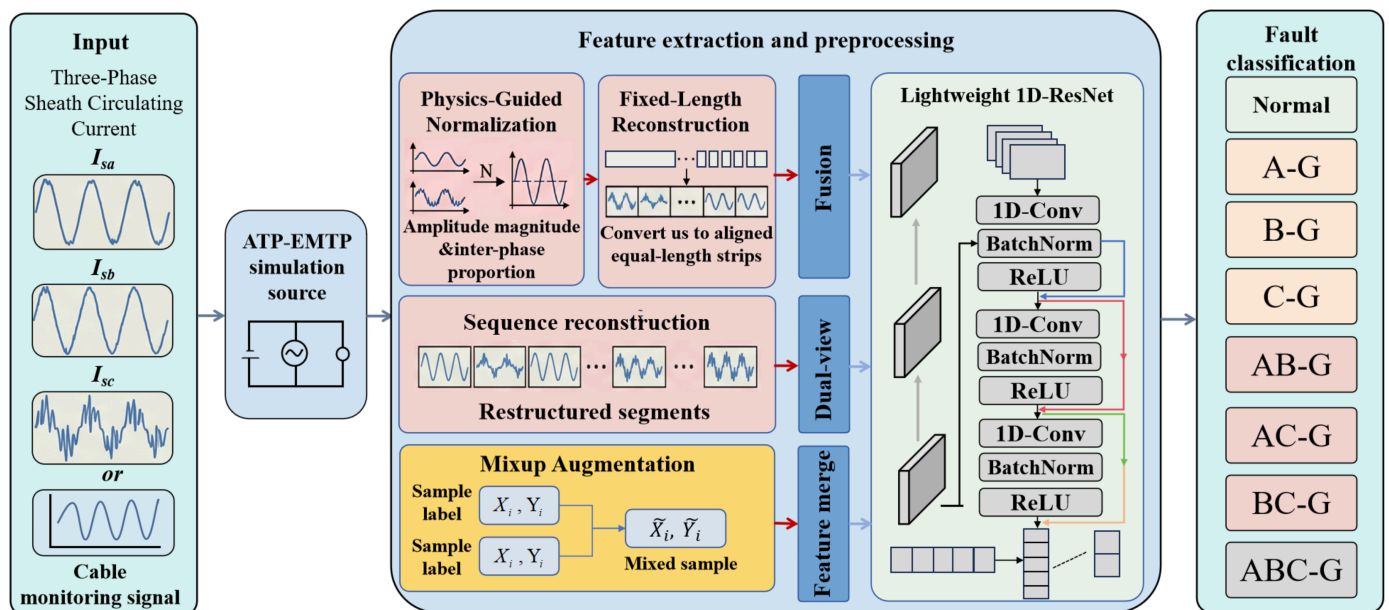


Figure 1. Overall framework of the proposed dual-view Mixup-ResNet method for intelligent monitoring and fault diagnosis of cable sheath circulating current signals.

2. Materials and Methods

2.1. ATP-EMTP-Based Cable Sheath Current Simulation Model

To construct a repeatable and controllable small-sample benchmark for fault diagnosis, an ATP-EMTP-based simulation model of cable sheath circulating currents was established using ATPDraw 7.7.0.13 and solved using the Alternative Transients Program (ATP), Salford 80386 translation, with the ATP source code dated 13 March 2019. Previous research has also demonstrated the feasibility of using ATP-EMTP to model sheath-current behavior and defect-related variations in cable systems [36]. The model was used to generate three-phase sheath current time-series signals under the normal operating condition as well as multiple fault conditions [11]. Focusing on a high-voltage XLPE cable and its sheath grounding loop, the simulation framework was designed to capture the dynamic variations in sheath currents under different grounding states, fault-phase combinations, and disturbance scenarios. Compared with static equivalent-parameter analysis, ATP-EMTP can represent electromagnetic transients and multi-phase coupling effects more faithfully [13,36], which makes it suitable for constructing time-series samples for intelligent monitoring-oriented fault diagnosis. The ATP-EMTP-based simulation platform used to generate the sheath circulating current samples is shown in Figure 2.

To improve the reproducibility of the ATP-EMTP simulation model, the main cable model parameters, grounding arrangement, fault configuration, and signal acquisition settings are summarized in Table 1. The simulated cable was a 110 kV single-core XLPE cable laid in parallel in soil. The line-core diameter was 48.6 mm, corresponding to an equivalent cross-sectional area of approximately 1855 mm². The sheath inner and outer diameters were 73 mm and 75.6 mm, respectively, and the phase spacing was 0.27 m. The model was constructed using the CABLE CONSTANTS module with a constant-parameter distributed cable representation. The simulated cable route consisted of four 125 m sections, giving a total length of 500 m. Grounding faults were modeled by connecting the faulted sheath phase(s) to ground through a low-resistance branch, while the non-faulted phase(s) remained under normal or high-resistance grounding states. The exported sheath current signals were sampled at 10 kHz, and the full 0.1 s record was used as one 1000-point input sample without sliding or overlapping windows. The three-phase sheath

currents were extracted from the measuring branches X0001A–X0003A, X0001B–X0003B, and X0001C–X0003C.

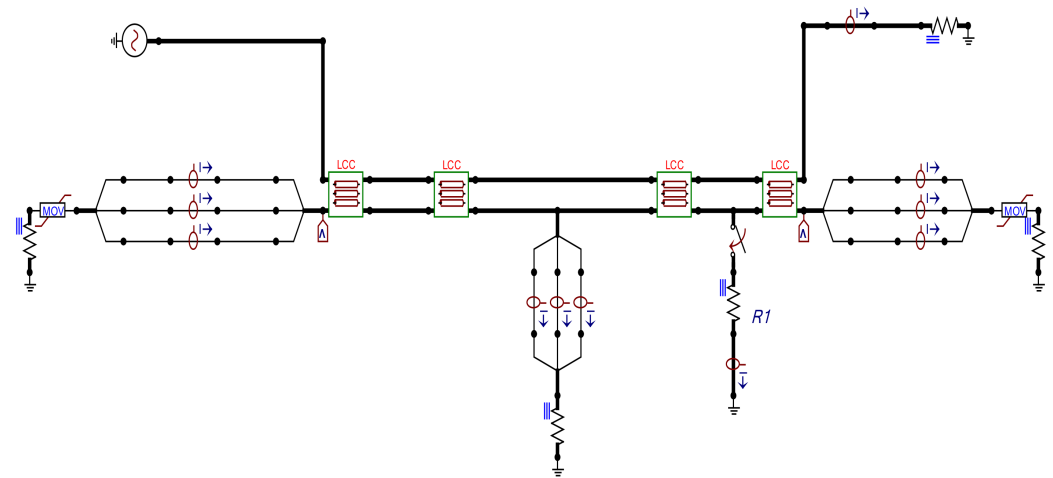


Figure 2. ATP-EMTP-based simulation model for cable sheath circulating current signal generation under normal and fault conditions.

Table 1. Main cable parameters, ATP-EMTP simulation configuration, and signal acquisition settings.

Item	Setting
Cable modeling module	CABLE CONSTANTS module
Cable representation	Constant-parameter distributed cable model
Voltage rating	110 kV
Cable type	Single-core XLPE cable
Installation and laying arrangement	Underground parallel laying
Number of single-core cables	3
Core diameter	48.6 mm
Equivalent core cross-sectional area	approximately 1855 mm ²
Sheath inner/outer diameter	73 mm/75.6 mm
Phase spacing	0.27 m
Cable section length	125 m
Number of cable sections	4
Total simulated cable length	500 m
Grounding/fault configuration	Faulted sheath phase(s) connected to ground through a low-resistance branch
Non-faulted phase state	Normal or high-resistance grounding state
Fault resistance	0.1 Ω
High-resistance grounding branch	1 k Ω
Fault inception time	0.05 s
Record length	0.1 s
ATP-EMTP simulation time step	1×10^{-7} s
Exported signal sampling interval	1×10^{-4} s
Sampling frequency	10 kHz
Sample length	1000 time points
Windowing method	Full 0.1 s record used as one sample; no sliding or overlapping window
Sheath current extraction branches	X0001A–X0003A, X0001B–X0003B, X0001C–X0003C
Saved current channels	CurrentA, CurrentB, CurrentC
CT/current measurement range	± 50 A
$I_{\max}$ for physical-range normalization	50 A

The objective of the simulation model is not limited to reproducing steady-state magnitude deviations; rather, it aims to characterize the overall temporal response of sheath

circulating currents during fault evolution. By configuring the normal operating state together with several representative abnormal conditions, the three-phase sheath current signals were extracted as raw observations. These signals form the basis for subsequent dual-view input construction, deep feature learning, and robustness assessment. In essence, the ATP-EMTP model provides a unified experimental platform for method validation, allowing different diagnostic strategies to be compared under the same fault scenarios and data source [9].

2.2. Fault Scenario Design and Dataset Construction

In accordance with the practical requirements of abnormal sheath current identification, eight operating conditions were considered in this study, namely the normal condition (Normal), single-phase grounding faults on phase A, phase B, and phase C (A-G, B-G, and C-G), and multi-phase composite faults (AB-G, AC-G, BC-G, and ABC-G). This fault set covers the normal state, single-phase faults, and composite multi-phase faults, and it is therefore able to reflect the variations in magnitude distribution and waveform morphology of sheath circulating current signals under different abnormal patterns [9,13].

The benchmark dataset used in the main experiment contained 40 labeled samples, with five samples for each operating condition. This small-sample setting was kept to reflect the limited availability of high-quality fault data in practical cable operation and maintenance.

For each operating condition, the five labeled samples were generated from independent ATP-EMTP simulation cases rather than by directly shifting, cropping, or repeatedly segmenting the same waveform. To make the sample-generation process more transparent, Table 2 reports the RMS current ranges calculated from the raw exported three-phase sheath-current waveforms of the five independent cases in each class before normalization. These RMS ranges provide a compact description of the actual current-response range covered by the exported waveforms in each operating condition.

Table 2. Operating conditions, sample distribution, and RMS current ranges of the independent ATP-EMTP cases.

Class	Samples	Phase-A RMS (A)	Phase-B RMS (A)	Phase-C RMS (A)	Faulted Branch
Normal	5	9.286–16.276	9.030–15.830	9.284–16.283	None
A-G	5	11.578–13.848	9.029–9.049	9.285–9.308	Phase A
B-G	5	9.287–9.306	10.401–11.402	9.284–9.308	Phase B
C-G	5	9.285–9.306	9.031–9.048	11.788–14.303	Phase C
AB-G	5	12.162–13.605	11.346–12.681	9.284–9.290	Phases A and B
AC-G	5	11.740–14.472	9.035–9.039	11.506–14.086	Phases A and C
BC-G	5	9.291–9.296	9.452–10.357	12.142–14.870	Phases B and C
ABC-G	5	12.005–13.594	10.486–11.428	12.975–15.015	Phases A, B and C

In the ATP-EMTP model, different fault categories were configured by activating the corresponding phase-to-ground fault branch. For the fault classes, the activated low-resistance grounding branch was fixed at $0.1\ \Omega$, while the high-resistance/non-faulted grounding branch was fixed at $1\ \text{k}\Omega$. Therefore, the grounding resistance was treated as a controlled simulation setting rather than a randomly varying variable. The class-dependent fault-related parameter was the activated faulted branch combination, as listed in Table 2, while the within-class variation was mainly introduced through different operating-current/loading cases.

The fault was introduced by closing the corresponding switch at 0.05 s within a 0.1 s simulation record. For the normal condition, no low-resistance fault branch was activated. Each full 0.1 s record was used as one sample, and no sliding, cropping, or overlapping

segmentation was used to enlarge the dataset. In this way, the samples used in different cross-validation folds were generated from independent ATP-EMTP cases and were not near-duplicate segments from the same simulation record.

In real power cable systems, it is difficult to collect a large number of measured cable fault samples. Full-current data under specific sheath grounding fault modes are even harder to obtain. This makes cable fault diagnosis a typical small-sample problem. In some cases, the number of samples in different fault classes may also be unbalanced.

For this reason, the dataset in this study was not designed to maximize the sample size. Instead, it was used to examine whether the proposed dual-view Mixup-ResNet method can work under limited data, strong inter-phase coupling, and multi-class fault conditions. The dataset construction, therefore, focused on covering the main fault classes, keeping the fault types representative, and ensuring that the experiment could be repeated.

In the main dataset, the grounding fault branch was modeled as a low-resistance grounding fault with $R_1 = 0.1 \Omega$. The $1 \text{ k}\Omega$ branch was used to represent the high-resistance/non-faulted grounding state in the simulation configuration. High-resistance grounding faults were not treated as an independent fault category in the present dataset and will be considered in future work.

2.3. Signal Preprocessing and Dual-View Input Representation

Cable sheath circulating current signals are typical electrical time-series data and exhibit non-stationarity and strong inter-phase coupling [8,10]. Different fault conditions are reflected not only in waveform shape variations but also in significant differences in absolute current magnitude and inter-phase proportional relationships [13]. These characteristics are directly related to fault severity and changes in conductive paths [9,11]. Therefore, when constructing an intelligent monitoring-oriented fault diagnosis model, it is necessary to preserve the physical magnitude information of the signals while enhancing the model's ability to capture waveform-shape differences [9].

Conventional z-score normalization can improve numerical consistency and facilitate model convergence. However, if it is used alone for sheath circulating current signals, the absolute amplitude information associated with fault severity, as well as the proportional relationships among the three-phase currents, may be weakened [8,10]. To address this issue, the proposed method first performs physical-range normalization on the three-phase signals according to the current transformer range and the allowable sheath current measurement range [9]. Let the original three-phase sheath current sequence be denoted as

$$\mathbf{X} = [\mathbf{x}_A, \mathbf{x}_B, \mathbf{x}_C] \in R^{T \times 3} \quad (1)$$

where T denotes the sequence length.

By linearly scaling the signal with a predefined physical upper bound $I_{\max}$, the physically normalized primary view is obtained as

$$\mathbf{X}_p = \frac{\mathbf{X}}{I_{\max}} \quad (2)$$

This operation unifies the numerical range while preserving the absolute magnitude differences among fault conditions and the proportional structure across phases, thereby improving the physical interpretability of the input representation.

In this study, $I_{\max}$ denotes the physical upper bound used for physical-range normalization, rather than the maximum value calculated from each individual sample. It was determined according to the current transformer measurement range and the maximum sheath circulating current considered in the ATP-EMTP simulation model. The equivalent CT/current measurement range used in the signal acquisition setting was

± 50 A, and $I_{\max}$ was set to 50 A in all experiments. Using a fixed physical upper bound avoids sample-dependent scaling and preserves the magnitude differences among different fault conditions.

To further strengthen the characterization of local waveform morphology and relative variation trends, a sample-wise auxiliary z-score view is constructed in addition to the physically normalized primary view. Specifically, the mean and standard deviation of each sample are computed on a per-phase basis, and the signal is standardized as

$$\mathbf{X}_z = \frac{\mathbf{X} - \mu(\mathbf{X})}{\sigma(\mathbf{X}) + \varepsilon} \quad (3)$$

where $\mu(\mathbf{X})$ and $\sigma(\mathbf{X})$ denote the sample-wise mean and standard deviation, respectively, and ε is a small constant introduced to avoid division by zero. Unlike approaches that rely on a single normalization strategy, the present study does not replace physical-range normalization with z-score normalization; instead, the two are treated as complementary views so that fault-intensity information and waveform-shape information can be retained simultaneously. The complete preprocessing procedure and the construction of the dual-view input tensor are illustrated in Figure 3.

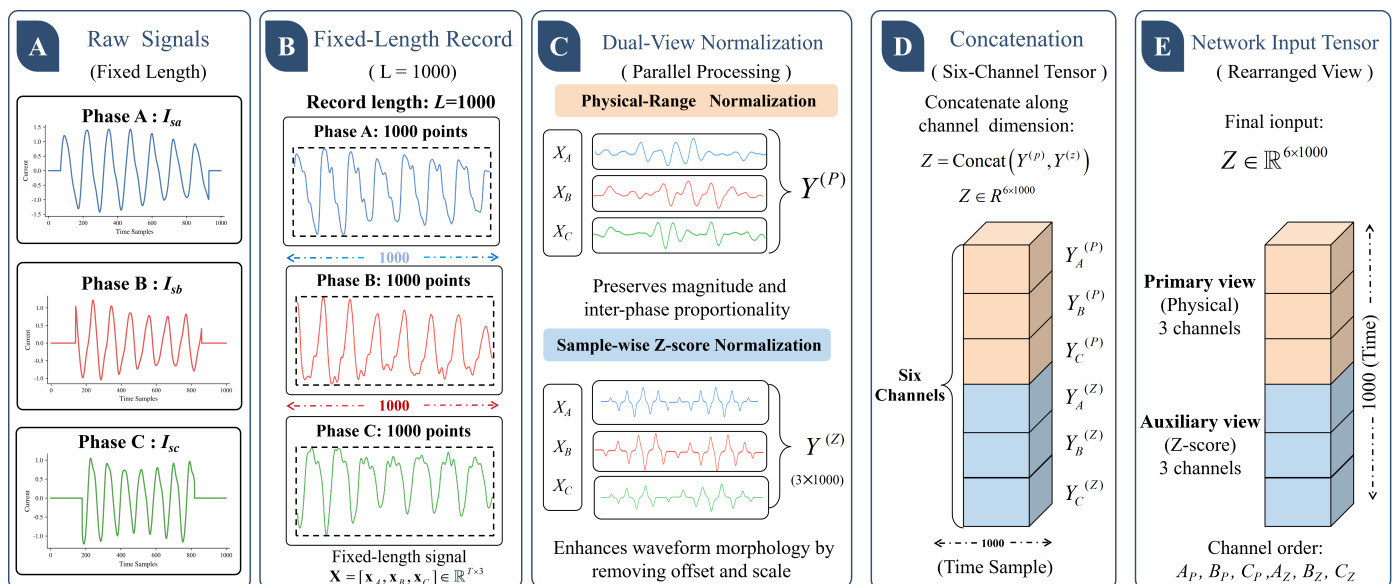


Figure 3. Signal preprocessing and dual-view input construction.

All exported ATP-EMTP simulation records have a fixed length of 1000 sampling points, and no trimming or zero-padding was applied. The sampling interval was 1×10^{-4} s, corresponding to a sampling frequency of 10 kHz, resulting in a total signal duration of 0.1 s per sample. This observation window covers the fault introduction at 0.05 s and the subsequent sheath-current response, providing sufficient temporal information for characterizing the fault-related waveform dynamics.

The physically normalized primary view $\mathbf{X}_p$ and the auxiliary z-score view $\mathbf{X}_z$ are then concatenated along the channel dimension to form the final input tensor,

$$\mathbf{X}_{in} = [\mathbf{X}_p; \mathbf{X}_z] \in \mathbb{R}^{T \times 6} \quad (4)$$

After dimension rearrangement, the resulting six-channel time-series tensor is fed into the subsequent one-dimensional residual network for end-to-end fault feature learning under small-sample conditions. The raw three-phase sheath current signals are transformed

into a physically normalized primary view and an auxiliary z-score view, which are then concatenated into a six-channel input tensor.

The proposed dual-view input representation combines physical priors with statistical morphological information. On the one hand, physical-range normalization preserves the magnitude information closely related to fault severity, enabling the model to better distinguish fault patterns with different levels of severity and different phase combinations. On the other hand, the auxiliary z-score view emphasizes local waveform variations and morphological differences, thereby improving the model's ability to characterize complex decision boundaries. By integrating these two complementary views, the proposed approach provides a more discriminative input representation for the lightweight deep network and enhances fault diagnosis performance and robustness in intelligent monitoring scenarios. To further illustrate the motivation for the proposed dual-view representation, representative three-phase sheath circulating current waveforms under typical operating conditions are shown in Figure 4.

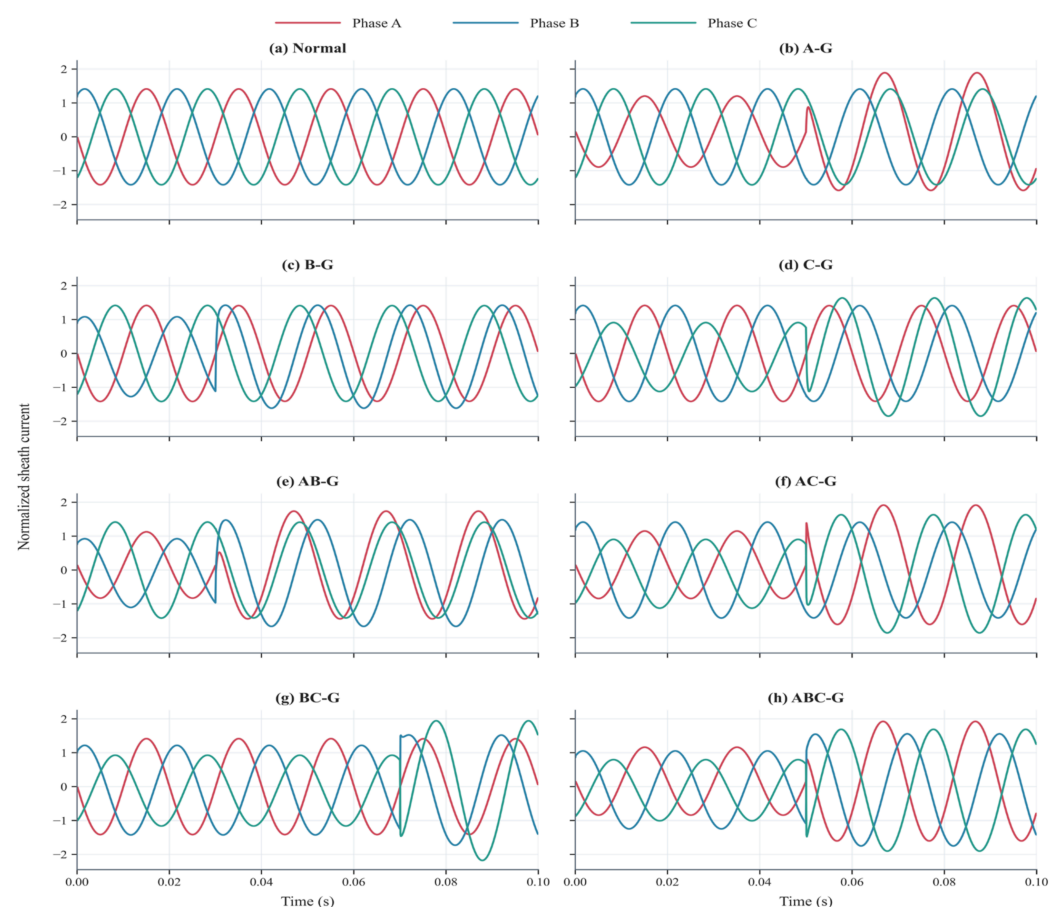


Figure 4. Representative three-phase sheath circulating current waveforms under different operating conditions: (a) normal operation; (b) phase-A grounding fault; (c) A-B composite grounding fault; (d) A-C composite grounding fault; (e) phase-B grounding fault; (f) phase-C grounding fault; (g) B-C composite grounding fault; (h) A-B-C composite grounding fault.

2.4. Mixup Data Augmentation and Robust Training Strategy

In cable sheath circulating current fault diagnosis, the number of available labeled samples is usually limited, and samples corresponding to composite faults or incipient abnormal states are particularly scarce [5,16]. Under such small-sample conditions, deep models tend to overfit local patterns in the training set, making the learned decision boundaries overly sensitive to perturbations and thereby weakening generalization under

unseen operating conditions and noisy environments [33]. To alleviate this issue, the proposed method introduces Mixup data augmentation during training and combines it with label smoothing, dropout, and cosine annealing to establish a robust training strategy for intelligent monitoring scenarios.

The core idea of Mixup is to linearly interpolate two training samples and their labels in the sample space, thereby generating virtual samples with smooth transitions between classes [35]. Let $\mathbf{x}_i$ and $\mathbf{x}_j$ denote two training samples, and $\mathbf{y}_i$ and $\mathbf{y}_j$ denote their corresponding labels. The generated Mixup sample can be written as

$$\begin{cases} \tilde{\mathbf{x}} = \lambda \mathbf{x}_i + (1 - \lambda) \mathbf{x}_j \\ \tilde{\mathbf{y}} = \lambda \mathbf{y}_i + (1 - \lambda) \mathbf{y}_j \end{cases} \quad (5)$$

where λ is the mixing coefficient, which is typically sampled from a Beta distribution [35],

$$\lambda \sim \text{Beta}(\alpha, \alpha) \quad (6)$$

In this study, the Mixup parameter α was set to 0.4. This value was adopted as a moderate interpolation setting according to the commonly used Mixup configuration and the physical characteristics of sheath circulating current signals. Since the input signals contain fault-related magnitude information and inter-phase proportional relationships, an excessively large Mixup coefficient may generate over-smoothed virtual samples with weaker physical interpretability. Therefore, $\alpha = 0.4$ was used in the main experiment as a conservative setting to provide regularization while avoiding overly strong interpolation. To further evaluate the influence of this parameter, a sensitivity analysis was conducted by varying α among 0, 0.1, 0.2, 0.4, 0.8, and 1.0.

It should be noted that the samples generated by Mixup are used only as virtual training samples for classifier regularization and decision-boundary smoothing. They are not regarded as physically simulated cable grounding-fault cases and are not used to define new operating conditions. A representative example of a Mixup-generated virtual training sample is shown in Figure 5. As illustrated, the mixed waveform is obtained by weighted interpolation of two original ATP-EMTP-generated samples under the same temporal support, while its role is limited to optimization-level regularization during training.

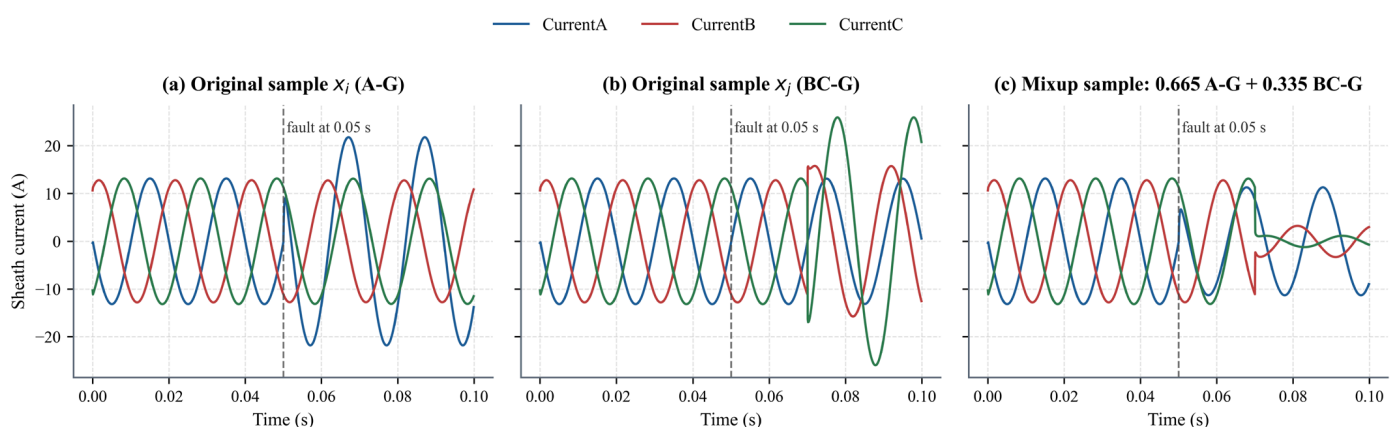


Figure 5. Illustration of a Mixup-generated virtual training sample. Panels (a) and (b) show two original three-phase sheath-current samples from the A-G and BC-G classes, respectively, and panel (c) shows their weighted Mixup combination. The generated waveform is used only as a virtual regularization sample during training and is not interpreted as a physically simulated cable fault case.

For sheath circulating current time-series signals, Mixup is particularly suitable for small-sample diagnosis tasks because different fault conditions often exhibit gradual

variations in magnitude distribution, inter-phase proportion, and local waveform structure [10,13]. On the one hand, constructing virtual samples through linear interpolation helps expand the coverage of the training distribution. On the other hand, compared with conventional augmentation methods such as noise injection, translation, or scaling, Mixup does not rely on manually designed local perturbation rules. Instead, it improves the model's ability to learn complex decision boundaries through input-space interpolation, making it more appropriate for multi-class and strongly coupled electrical fault recognition [35].

On top of Mixup, label smoothing is further introduced to reduce excessive confidence in hard labels. Under small-sample conditions, the standard cross-entropy loss tends to drive the model toward overly sharp posterior distributions. By slightly softening the supervisory targets, label smoothing can improve probability calibration and enhance generalization [37]. In the present study, cross-entropy loss with label smoothing is employed, and the smoothing factor is set to 0.05. Accordingly, the final training objective can be written as

$$L = - \sum_{k=1}^K \tilde{y}_k^{(ls)} \log p_k \quad (7)$$

$$\tilde{y}_k^{(ls)} = (1 - \varepsilon_{ls}) \tilde{y}_k + \frac{\varepsilon_{ls}}{K} \quad (8)$$

where $\tilde{y}_k^{(ls)}$ denotes the smoothed mixed label for class k , $\tilde{y}_k$ is the mixed target generated by Mixup, ε_{ls} is the label smoothing factor, and K is the total number of classes.

The learning rate is updated using cosine annealing [38] according to

$$\eta_t = \eta_{\min} + \frac{1}{2}(\eta_0 - \eta_{\min}) \left(1 + \cos \frac{t\pi}{T} \right) \quad (9)$$

where η_t is the learning rate at epoch t , η_0 is the initial learning rate, $\eta_{\min}$ is the minimum learning rate, and T denotes the total number of training epochs.

In addition, dropout is inserted before the final classification layer to further reduce the risk of overfitting during training, with the dropout rate set to 0.2 [39]. This strategy randomly suppresses a portion of feature responses during optimization, thereby weakening co-adaptation among hidden representations and promoting more robust feature learning. The model is optimized using Adam [40], with an initial learning rate of 5×10^{-4} and a weight decay coefficient of 1×10^{-4} . For learning rate scheduling, cosine annealing is adopted so that the learning rate decays smoothly during training, balancing rapid convergence in the early stage and refined optimization in the later stage [38]. The batch size is set to 16, and the total number of training epochs is 200.

Overall, the performance improvement of the proposed method does not rely on Mixup alone. Instead, it is achieved through a joint training strategy composed of Mixup augmentation, label smoothing, dropout, and cosine annealing. Together with the dual-view input representation, this training mechanism constitutes an essential part of the proposed framework and provides strong support for stable, efficient, and generalizable fault diagnosis using the subsequent one-dimensional residual network.

2.5. Lightweight 1D-ResNet Architecture

For online intelligent monitoring of cable sheath circulating current signals, the fault diagnosis model is expected not only to achieve high classification accuracy but also to maintain computational efficiency and deployment feasibility [16]. Since sheath circulating current signals are inherently multi-channel time-series data, converting them into two-dimensional images may introduce additional preprocessing overhead and redundant features [32]. Therefore, this study adopts a lightweight one-dimensional residual network

(1D-ResNet) as the core diagnostic model to perform end-to-end feature learning directly on the time-series input [34].

Recent studies have also shown that one-dimensional neural networks are effective for directly processing raw one-dimensional sensor signals. For example, 1D-CNN has been used for one-dimensional signal recognition tasks without image transformation [41], and 1D-ResNet has been applied to multivariate time-series fault detection using raw sensor data [42]. These studies further support the use of one-dimensional architectures for current-signal diagnosis, because they avoid the additional preprocessing cost and redundant feature construction associated with 1D-to-2D image conversion.

After preprocessing, each sample is represented as a six-channel sequence tensor with a length of 1000, consisting of the three-phase physically normalized primary view and the three-phase auxiliary z-score view. This input format preserves the absolute magnitude information and inter-phase proportional relationships associated with fault severity, while also enhancing the representation of local waveform variations and morphological differences. As a result, it provides a more discriminative input basis for subsequent deep feature extraction.

In terms of network architecture, the model first employs a one-dimensional convolutional stem for initial feature extraction. This stem contains a 1D convolution layer with a kernel size of 7 and a stride of 2, projecting the input channels into a 16-dimensional feature space, followed by batch normalization and a ReLU activation function to improve feature representation. Zero padding with a padding size of 3 was applied in the convolutional stem. With a kernel size of 7 and a stride of 2, this setting reduces the sequence length from 1000 to 500, producing an output size of 16×500 .

A max-pooling layer is then applied to reduce the temporal dimension, compress redundant information, and improve the efficiency of the subsequent feature extraction process. Compared with shallow convolutional structures, this design is able to capture local temporal patterns effectively while maintaining relatively low computational complexity [34].

In the backbone, three residual stages are constructed, with the number of channels progressively expanded to 16, 32 and 64. Each stage contains two basic residual blocks, and each block consists of two one-dimensional convolution layers, batch normalization, and ReLU activation. Residual connections are implemented through identity mapping or linear projection. The residual mapping of a basic unit can be expressed as [34]

$$\mathbf{y} = \mathcal{F}(\mathbf{x}, \{W_i\}) + \mathbf{x} \quad (10)$$

where $\mathbf{x}$ and $\mathbf{y}$ denote the input and output feature maps, respectively, and $\mathcal{F}(\mathbf{x}, \{W_i\})$ represents the residual transformation parameterized by trainable weights $\{W_i\}$.

When the input and output dimensions are inconsistent or temporal downsampling is required, a 1×1 convolution is introduced in the shortcut branch for dimensional matching. This residual design effectively alleviates the degradation problem encountered in deeper networks and enables stable hierarchical feature learning with a limited parameter budget [34], making it well-suited to small-sample fault diagnosis tasks.

After residual feature extraction, adaptive global average pooling is employed to compress the temporal features into a fixed-length representation vector, thereby reducing the number of parameters and improving tolerance to slight variations in sequence length. In the classification head, a dropout layer with a rate of 0.2 is inserted before the final fully connected layer to mitigate the risk of overfitting under small-sample training conditions [39]. The final classification layer outputs the predicted label among the eight operating conditions, enabling unified identification of the normal condition, single-phase grounding faults, and multi-phase composite faults.

Compared with a basic CNN, the adopted 1D-ResNet enhances deep feature extraction through residual connections and is more capable of capturing multiscale and progressive fault patterns in sheath circulating current signals [34]. Compared with larger and more complex architectures, the proposed model remains compact and lightweight, which is favorable for deployment on resource-constrained online monitoring terminals. More importantly, the performance gain of the proposed approach does not stem from network depth alone. Rather, it is achieved through the coordinated design of dual-view input representation, lightweight residual feature extraction, and robust training strategies, allowing the model to achieve high diagnostic accuracy while maintaining competitive disturbance tolerance under small-sample conditions.

Accordingly, the lightweight 1D-ResNet developed in this study not only satisfies the accuracy requirements of cable sheath circulating current fault diagnosis but also provides a lightweight algorithmic basis for future online intelligent monitoring and early fault-warning studies after further validation with measured data. The detailed configuration of the proposed lightweight 1D-ResNet is summarized in Table 3.

Table 3. Configuration of the proposed lightweight 1D-ResNet.

Stage	Operator	Kernel/Stride/Padding	Channels	Output Size
Input	Dual-view current tensor	—	6	6×1000
Stem	Conv1D + BN + ReLU	7/2/3	16	16×500
Pool	MaxPool1D	3/2	16	16×250
Stage 1	2 residual blocks	3/1	16	16×250
Stage 2	2 residual blocks	3/2	32	32×125
Stage 3	2 residual blocks	3/2	64	64×63
Head	Adaptive AvgPool1D	—	64	64×1
Regularization	Dropout	$p = 0.2$	64	64×1
Classifier	Fully connected	—	8	8

2.6. Evaluation Protocol and Metrics

To objectively evaluate the generalization capability of the proposed method under small-sample conditions, stratified five-fold cross-validation was adopted. Specifically, the 40 labeled samples were divided into five mutually exclusive subsets while preserving the class distribution in each fold. Since each class contains five samples, every fold uses one sample from each class for testing and the remaining four samples for training. Because the samples in each class were generated from independent simulation cases, no waveform segments originating from the same simulation record were simultaneously assigned to the training and testing subsets. The final performance is reported as the average over the five folds, and the corresponding standard deviation is used to characterize model stability. To reduce the influence of a single data partition under the limited-sample setting, a unified 5×5 repeated stratified cross-validation protocol was used for the final performance estimation, consisting of five repetitions of stratified five-fold cross-validation. The mean value, standard deviation, and 95% confidence interval were reported.

The main evaluation metrics include diagnostic accuracy, weighted F1-score, and the normalized confusion matrix. Let N denote the total number of test samples, and let y_i and $\hat{y}_i$ represent the true and predicted labels of the i -th sample, respectively. Accuracy is defined as

$$\text{Accuracy} = \frac{1}{N} \sum_{i=1}^N \mathbb{I}(\hat{y}_i = y_i) \quad (11)$$

where $\mathbb{I}(\cdot)$ is the indicator function. For the k -th class, precision, recall, and F1-score are given by

$$\text{Precision}_k = \frac{TP_k}{TP_k + FP_k} \quad (12)$$

$$\text{Recall}_k = \frac{TP_k}{TP_k + FN_k} \quad (13)$$

$$F1_k = \frac{2 \cdot \text{Precision}_k \cdot \text{Recall}_k}{\text{Precision}_k + \text{Recall}_k} \quad (14)$$

In the main experiment, the overall classification performance is measured using the weighted F1-score, defined as

$$F1_{\text{weighted}} = \sum_{k=1}^K \frac{n_k}{N} F1_k \quad (15)$$

where n_k is the number of samples in class k , and K is the total number of classes.

To further assess robustness under disturbed conditions, relative Gaussian noise was injected into the input signals in the robustness study. Let $\mathbf{x}$ denote the original input, $\sigma_{\mathbf{x}}$ its sample-wise standard deviation, and r the noise ratio. The noisy input can then be written as

$$\mathbf{x}_{\text{noisy}} = \mathbf{x} + \boldsymbol{\eta} \quad \boldsymbol{\eta} \sim \mathcal{N}(0, (r\sigma_{\mathbf{x}})^2) \quad (16)$$

In this study, $r = 0.12$ was adopted in the robustness experiments, and the performance of different model configurations was compared on both clean and noisy test sets.

3. Results

3.1. Implementation Details

During training, the Adam optimizer [40] was adopted with an initial learning rate of 5×10^{-4} , a weight decay coefficient of 1×10^{-4} , a batch size of 16, and a total of 200 training epochs. For the main model, the Mixup parameter α was set to 0.4 [35], the label smoothing factor was set to 0.05 [37], and the dropout rate before the classification layer was set to 0.2 [39]. The Mixup coefficient was further examined through sensitivity analysis, while the label smoothing factor of 0.05 and the dropout rate of 0.2 were selected as moderate regularization settings to reduce overconfidence and overfitting under the small-sample condition, rather than through exhaustive grid search. Cosine annealing [38] was employed for dynamic learning rate scheduling. Together with the dual-view input representation, these settings constitute the complete implementation of the proposed method.

All deep-learning experiments were implemented in Python 3.10.19 using PyTorch 2.9.1+cu128 on a Windows-based workstation equipped with an Intel Core i9 CPU and an NVIDIA GeForce RTX 5060 Laptop GPU. CUDA 12.8 was available for GPU acceleration.

A unified 5×5 repeated stratified cross-validation protocol was used for the final performance estimation of the proposed model. The diagnostic performance was mainly evaluated in terms of accuracy, weighted F1-score, macro F1-score, and the confusion matrix. The main training hyperparameters and evaluation settings are summarized in Table 4.

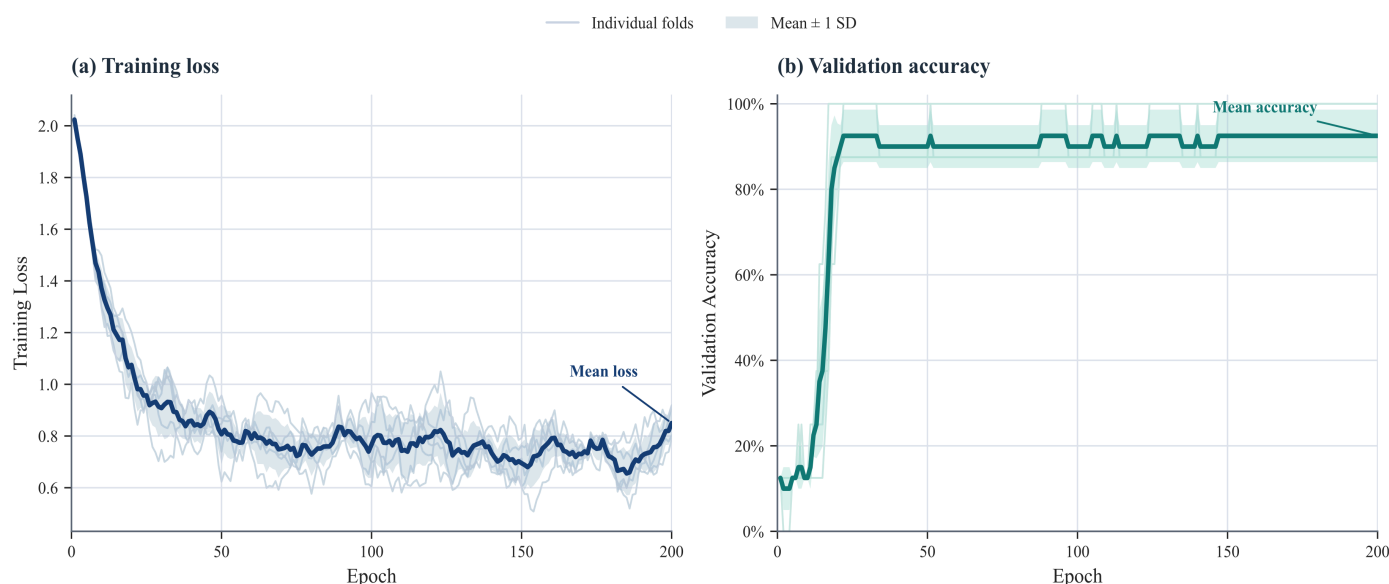
In the main experiment, the final performance of the proposed method was evaluated using the unified 5×5 repeated stratified cross-validation protocol. For the confusion-matrix visualization, a representative stratified five-fold result was used to illustrate the class-wise recognition behavior. For the ablation and robustness studies, mean accuracy, standard deviation, and macro F1-score were reported under the corresponding component-level evaluation settings to analyze the contribution of each module.

Table 4. Main training hyperparameters and evaluation settings.

Item	Setting
Optimizer	Adam
Initial learning rate	5×10^{-4}
Weight decay	1×10^{-4}
Batch size	16
Epochs	200
Cross-validation	5×5 repeated stratified cross-validation for final performance estimation
Mixup coefficient α	0.4
Label smoothing	0.05
Dropout rate	0.2
Learning rate schedule	Cosine annealing
Input representation	Dual-view current signals
Input channels	6
Sequence length	1000 sampling points
Robustness setting	12% relative Gaussian noise

3.2. Convergence Analysis and Overall Diagnostic Performance

Figure 6 presents the mean training loss and mean validation accuracy curves of the proposed dual-view Mixup-ResNet over five-fold cross-validation, where the shaded regions indicate the variation across folds. It can be observed that the mean training loss decreases rapidly during the early stage of training and gradually stabilizes thereafter. At the same time, the mean validation accuracy rises quickly and remains at a high level throughout most of the training process, indicating that the proposed method can be optimized in a stable manner even under small-sample conditions. The coordinated effect of the physically normalized primary view, the auxiliary z-score view, Mixup augmentation, and cosine annealing prevents persistent oscillation and severe overfitting during training.

**Figure 6.** Mean training loss and mean validation accuracy curves under the cross-validation setting.

To reduce the influence of a single data partition under the limited-sample setting and to ensure consistency with the Mixup sensitivity analysis, the main diagnostic performance was rechecked under a unified 5×5 repeated stratified cross-validation protocol. The summary diagnostic performance is listed in Table 5. As shown in Table 5, the proposed method with $\alpha = 0.4$ achieved an average accuracy of 91.50%, with a standard deviation

of 1.37% and a 95% confidence interval of 90.30–92.70%. The weighted F1-score reached 91.69%, with a standard deviation of 1.19% and a 95% confidence interval of 90.65–92.73%. The macro F1-score was 91.69%, with a standard deviation of 1.19% and a 95% confidence interval of 90.65–92.73%.

Table 5. Summary of the diagnostic performance of the proposed method under the unified 5×5 repeated stratified cross-validation protocol.

Metric	Mean	Std.	95% CI
Accuracy	91.50%	1.37%	90.30–92.70%
Weighted F1	91.69%	1.19%	90.65–92.73%
Macro F1	91.69%	1.19%	90.65–92.73%

These results provide a more conservative and internally consistent estimate of the proposed model under repeated resampling. Compared with reporting only one five-fold partition or an earlier single-run result, the unified 5×5 repeated stratified cross-validation protocol provides a more transparent estimate of model stability under the current small-sample setting. Nevertheless, because the dataset contains only 40 labeled samples, the reported results should still be interpreted as preliminary feasibility evidence under a small-sample simulation setting rather than a definitive estimate of field performance.

3.3. Confusion Matrix and Class-Wise Performance Analysis

To further examine the class-wise recognition capability of the proposed method, a normalized confusion matrix was constructed from the aggregated predictions of the five-fold cross-validation, as shown in Figure 7. In the confusion matrix, the horizontal axis represents the predicted class, the vertical axis represents the true class, and the diagonal elements correspond to the recall of each class. As illustrated in Figure 7, the proposed method achieves high recognition performance for most operating conditions. In particular, the recall values of the Normal, A-G, AB-G, AC-G, BC-G, and ABC-G classes all reach 100%, indicating that the model can reliably distinguish the normal condition from several composite fault conditions. The class-wise diagnostic performance of the proposed method is visualized by the normalized confusion matrix in Figure 7.

By comparison, the recall values of the B-G and C-G single-phase grounding faults are both 80%. The remaining confusion between B-G and C-G can be explained by the similarity of their electrical response characteristics under some weak single-phase grounding conditions. First, their sheath-current magnitudes may partially overlap when the grounding/fault resistance and operating current level lead to close current amplitudes. Second, because the three-phase sheath currents are coupled through the cable sheath loop and grounding branches, a single-phase grounding fault can also induce correlated changes in the non-faulted phases. This inter-phase coupling may reduce the distinction between B-G and C-G in the physically normalized view. Third, the local waveform shapes of B-G and C-G may remain similar within the 0.1 s observation window, especially when the transient distortion after fault inception is weak. Therefore, the misclassification is mainly attributed to the overlap of current magnitude, phase-coupling response, and local waveform morphology under limited-sample conditions.

A small number of B-G samples are misclassified as Normal, whereas a small number of C-G samples are misclassified as ABC-G. This suggests that, under the present small-sample setting, certain single-phase grounding faults still share partially overlapping characteristics with adjacent classes, especially when the fault magnitude is weak or the class boundary is less pronounced. Nevertheless, the confusion matrix still shows that the model can distinguish most operating conditions under the limited-sample setting. The

overall performance under the unified 5×5 repeated stratified cross-validation protocol is reported consistently in Table 5.

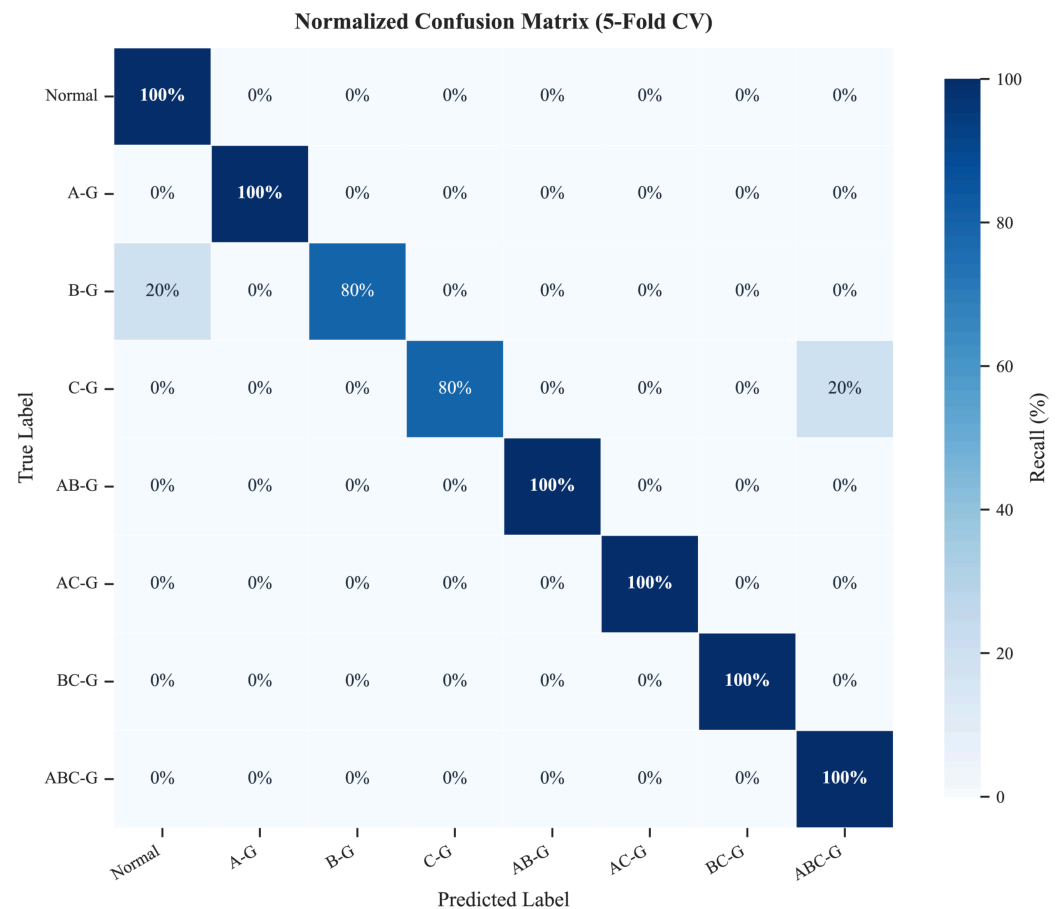


Figure 7. Normalized confusion matrix of the proposed dual-view Mixup-ResNet.

3.4. Comparison with Reference Methods

To validate the effectiveness of the proposed method in multi-class fault diagnosis of cable sheath circulating currents, several representative reference methods were selected for comparison, including classical feature-based models and standard end-to-end deep learning baselines [16,25]. Representative CNN- and LSTM-based approaches have also been reported in related cable diagnosis studies [10,33]. In general, conventional shallow learning methods rely heavily on handcrafted or low-level statistical features [25,28]. As a result, they may struggle to capture the deep fault patterns embedded in sheath circulating current signals, which are non-stationary, strongly coupled across phases, and rich in temporal structure [8,10]. Deep learning methods, by contrast, can automatically learn more discriminative representations from raw time-series data [16,32]. However, without a suitable input representation and robust training strategy, they may still suffer from overfitting and poor generalization in small-sample settings [16,34]. Therefore, SVM, KNN, and Random Forest were adopted as classical feature-based baselines, while plain 1D-CNN and plain 1D-ResNet were used as standard end-to-end time-series baselines. All reference methods were evaluated using the same dataset partitioning and cross-validation protocol as the proposed method.

From the overall comparison results in Table 6, the classical feature-based models show limited performance, with Random Forest achieving the best result among them at an accuracy of 83.4% and a weighted F1-score of 0.818. The plain 1D-CNN achieved an accuracy of 84.0%, while the plain 1D-ResNet achieved an accuracy of 85.0%, indicating

that simply introducing an end-to-end convolutional or residual structure provides only limited improvement under the current small-sample setting.

Table 6. Performance comparison of different diagnostic methods for cable sheath circulating current fault diagnosis.

Method	Feature/Input Type	Accuracy (%)	Weighted F1-Score
SVM	Statistical features	70.0	0.685
KNN (k = 3)	Statistical features	65.8	0.652
Random Forest	Statistical features	83.4	0.818
Plain 1D-CNN	Time-series input	84.0	0.8359
Plain 1D-ResNet	Time-series input	85.0	0.8540
Dual-View Mixup-ResNet	Dual-view current signals	91.5	0.9169

Note: Bold values indicate the best-performing method or setting in the corresponding comparison.

By contrast, the proposed dual-view Mixup-ResNet achieved the highest accuracy of 91.5% and the highest weighted F1-score of 0.9169 under the unified repeated cross-validation protocol. This result indicates that the improvement is not obtained by the residual backbone alone, but by the coordinated use of physical-range normalization, dual-view input representation, residual feature extraction, and Mixup-based regularization.

To provide a more intuitive explanation of the superiority of the proposed representation, the feature distributions learned by different settings are visualized in Figure 8.

To further evaluate the feasibility of the proposed model for online monitoring applications, the computational complexity of different diagnostic models was compared in terms of parameter number, FLOPs, and single-sample inference time, as shown in Table 7. The parameter number and FLOPs were calculated using the corresponding model input size, and the inference time was measured under evaluation mode with batch size 1 after warm-up runs. For Feature-SVM, Params and FLOPs are marked as ‘—’ because their computational cost mainly depends on handcrafted feature extraction and non-neural inference procedures.

Table 7. Computational complexity comparison of representative diagnostic models.

Method	Input	Params (M)	FLOPs (M)	Inference Time (ms/Sample)
Feature-SVM	75 handcrafted features	—	—	4.38
Plain 1D-CNN	3×1000	0.0098	6.11	1.41
Plain 1D-ResNet	3×1000	0.0614	10.37	6.31
Proposed Dual-View Mixup-ResNet	6×1000	0.0617	10.70	6.65

Note: Bold values indicate the best-performing method or setting in the corresponding comparison.

As shown in Table 7, the proposed Dual-View Mixup-ResNet contains only 0.0617 M parameters and requires 10.70 M FLOPs, with an inference time of 6.65 ms per sample. Compared with the plain 1D-ResNet, the proposed model introduces only a slight increase in computational cost while achieving higher diagnostic accuracy. This indicates that the performance improvement is obtained without a substantial increase in computational burden, supporting its potential feasibility for future online cable condition monitoring after further validation on measured data.

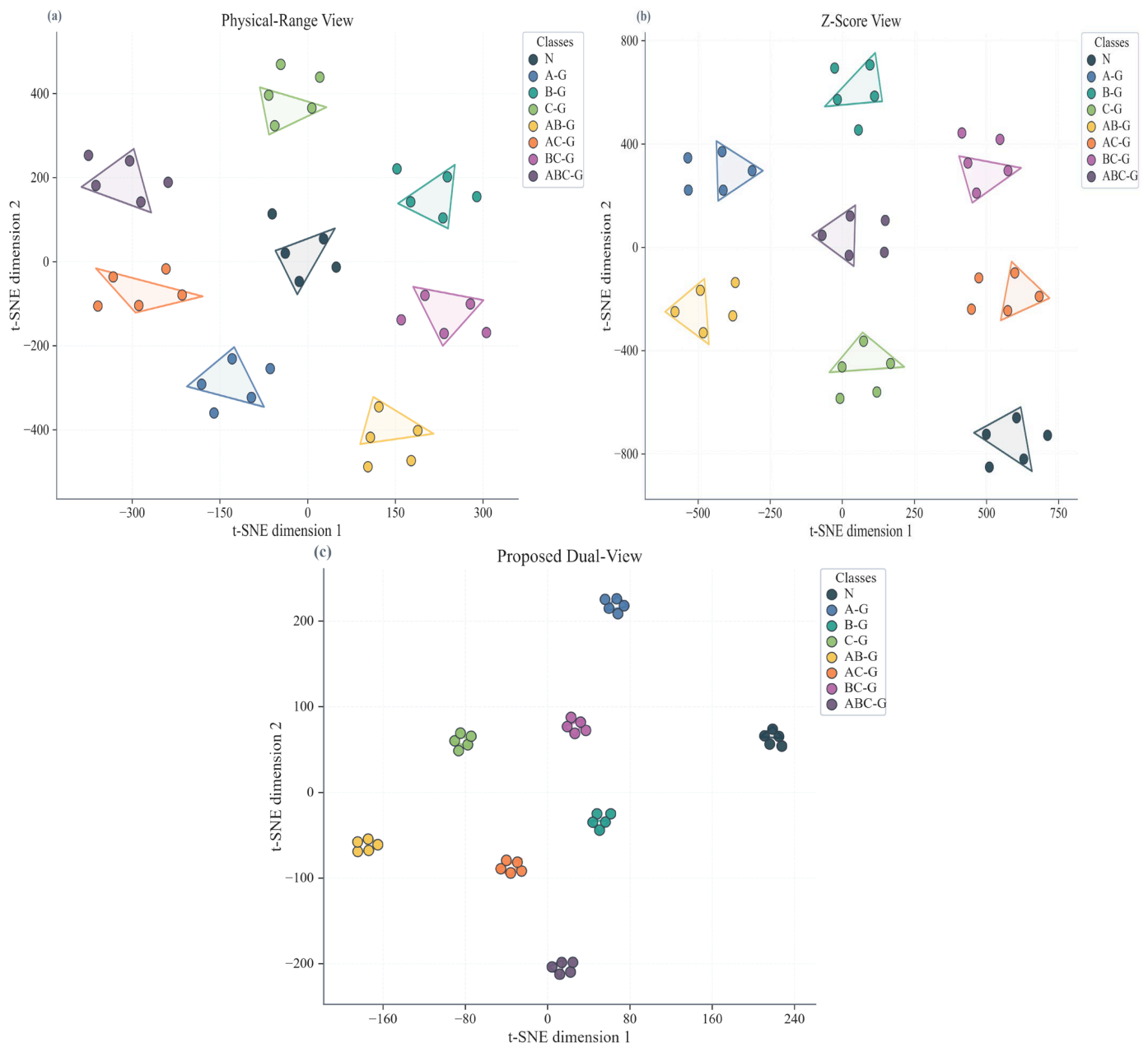


Figure 8. Feature-space visualization of the learned features under different input representations: (a) physical-range view; (b) z-score view; (c) proposed dual-view representation.

3.5. Ablation and Robustness Analysis

To further clarify the contribution of the major components, systematic ablation experiments were carried out with respect to the network backbone, input representation, and training strategy. To keep the ablation results consistent with the revised main performance evaluation, all model variants were evaluated under the same unified 5×5 repeated stratified cross-validation protocol, and the results are summarized in Table 8.

As shown in Table 8, the CNN baseline achieved an average accuracy of 84.00%, indicating that a shallow convolutional classifier has limited ability to capture the strongly coupled temporal patterns in sheath circulating current signals. After replacing the CNN baseline with a plain residual structure, the average accuracy reached 85.00%, showing only a limited improvement over the CNN baseline. This indicates that the residual backbone alone is not sufficient to explain the final performance gain under the current small-sample setting.

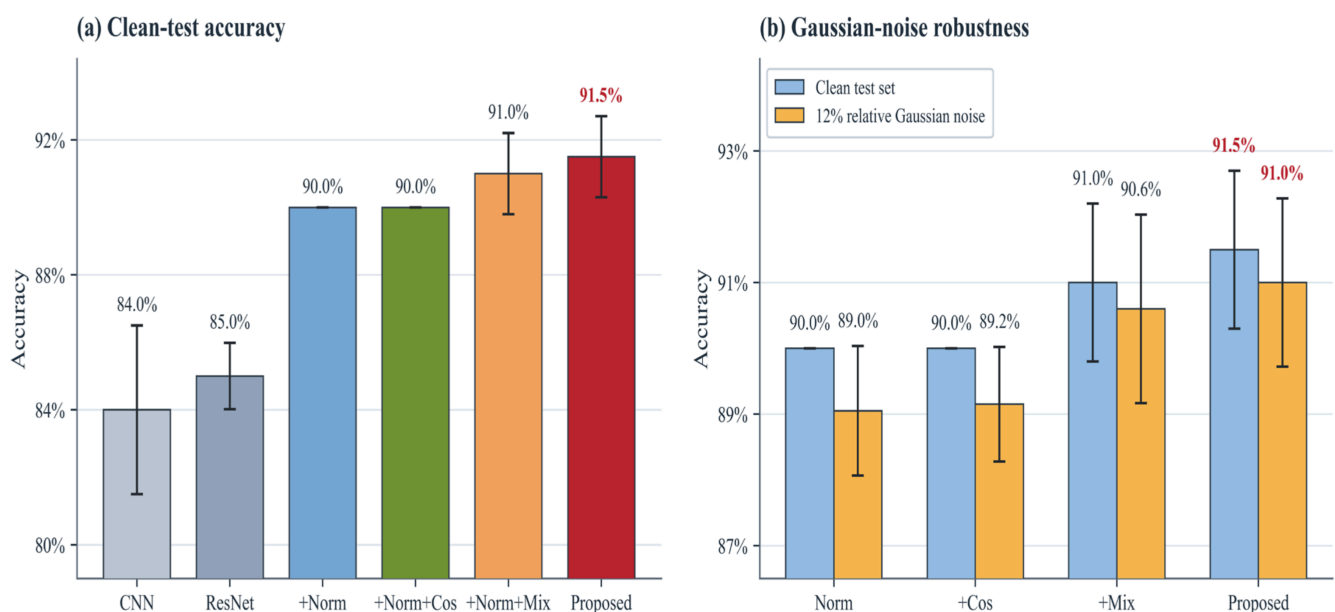
Table 8. Ablation results of different model variants under the unified 5×5 repeated stratified cross-validation protocol.

Variant	Accuracy (%)	95% CI (%)	Weighted F1	Macro F1
CNN Baseline	84.00 ± 2.85	81.50–86.50	0.8359	0.8359
ResNet	85.00 ± 1.12	84.02–85.98	0.8540	0.8540
ResNet + Physical Norm	90.00 ± 0.00	90.00–90.00	0.9027	0.9027
ResNet + Physical Norm + Cosine	90.00 ± 0.00	90.00–90.00	0.9027	0.9027
ResNet + Physical Norm + Mixup	91.00 ± 1.37	89.80–92.20	0.9133	0.9133
Proposed dual-view Mixup-ResNet	91.50 ± 1.37	90.30–92.70	0.9169	0.9169

When physical-range normalization was introduced, the average accuracy increased to 90.00%, suggesting that preserving magnitude-related fault information provides a more evident contribution than using the residual backbone alone. Adding cosine annealing did not further change the average accuracy in this setting, while introducing Mixup improved the average accuracy to 91.00%, with a weighted F1-score of 0.9133.

The complete dual-view Mixup-ResNet achieved the best overall performance among the evaluated variants, with an average accuracy of 91.50% and a weighted F1-score and macro F1-score of 0.9169. This result is consistent with the main performance reported in Table 5 and indicates that the final performance gain mainly comes from the coordinated contribution of physical-range normalization, dual-view input representation, Mixup-based regularization, and the lightweight residual classifier, rather than from the residual backbone alone.

Because the dataset is class-balanced, the weighted F1-score and macro F1-score are very close in this ablation analysis. The zero standard deviation observed for some variants reflects that the same average accuracy was obtained across the repeated evaluations under the current small dataset, rather than indicating that the model has no uncertainty in practical applications. For a more intuitive visualization of the ablation and noise-test results, Figure 9 presents the clean-test accuracy of different variants and the comparison between clean-test and noisy-test accuracy under 12% relative Gaussian noise.

**Figure 9.** Ablation and Gaussian-noise robustness analysis of different model variants: (a) clean-test accuracy under the unified 5×5 repeated stratified cross-validation protocol; (b) comparison of clean-test and noisy-test accuracy under 12% relative Gaussian noise.

In the Gaussian-noise robustness study, the clean-test and noisy-test accuracies of several representative variants were compared under 12% relative Gaussian noise, as shown in Figure 9b. The Physical Norm variant achieved 90.0% on the clean test set and 89.0% under Gaussian noise, while the Physical Norm + Cosine variant achieved 90.0% and 89.2%, respectively. After Mixup was introduced, the clean-test accuracy increased to 91.0%, and the noisy-test accuracy reached 90.6%. The proposed dual-view Mixup-ResNet achieved the highest clean-test accuracy of 91.5% and maintained a noisy-test accuracy of 91.0%.

These results indicate that the proposed method maintains competitive Gaussian-noise tolerance under the tested 12% relative Gaussian noise condition. However, this robustness observation is limited to the tested Gaussian-noise setting and should not be interpreted as evidence of practical robustness against broader field disturbances. Further evaluation under more realistic disturbances, such as sensor bias, missing data, harmonic interference, inverter-driven switching noise, and load fluctuation, is still needed.

3.6. Sensitivity Analysis of Mixup Coefficient

A sensitivity analysis was conducted to examine the influence of the Mixup coefficient α under the same unified 5×5 repeated stratified cross-validation protocol used for the revised main performance evaluation. As shown in Table 9, when α was set to 0.4, the proposed method achieved an average accuracy of 91.50% and a weighted F1-score of 91.69%. This result is consistent with the main performance reported in Table 5, where the macro F1-score was also 91.69%. Therefore, the main experiment and the Mixup sensitivity analysis now follow the same evaluation protocol and report consistent values.

Table 9. Sensitivity analysis of the Mixup coefficient α under the unified 5×5 repeated stratified cross-validation protocol.

Mixup Coefficient α	Accuracy (%)	Weighted F1-Score (%)
0	93.00 ± 2.74	93.05 ± 2.72
0.1	93.00 ± 2.09	92.99 ± 2.07
0.2	92.50 ± 0.00	92.53 ± 0.08
0.4	91.50 ± 1.37	91.69 ± 1.19
0.8	93.00 ± 1.12	93.09 ± 1.04
1.0	91.50 ± 1.37	91.59 ± 1.32

The diagnostic performance varied within a relatively narrow range when α changed from 0 to 1.0, indicating that the proposed framework is not strongly dependent on a single Mixup coefficient. Although $\alpha = 0.8$ achieved slightly higher average accuracy and weighted F1-score in the sensitivity analysis, $\alpha = 0.4$ was retained as a moderate interpolation setting because it provides regularization while avoiding overly strong interpolation of physically meaningful sheath-current signals. When α was further increased to 1.0, the performance decreased, suggesting that overly strong interpolation may over-smooth the virtual samples and weaken class-discriminative information. Therefore, Mixup is regarded as a regularization strategy for improving decision-boundary smoothness, rather than a method for generating physically real fault samples.

4. Discussion

4.1. Effectiveness of the Dual-View Representation

The experimental results show that the dual-view representation method proposed in this paper can effectively improve the fault diagnosis performance of the cable sheath loop signal. Unlike the traditional signal characterization method that adopts a single

normalization strategy, this method retains two types of differentiated and complementary diagnostic feature information through a dual-view input structure. Among them, the physical-range normalization view can fully retain the characteristics associated with the absolute amplitude of the current and the proportion between the three phases, which are highly related to the severity of the fault, the grounding method and the coupling characteristics between the phases; and in the fault diagnosis of the sheath loop, various grounding faults will not only change the signal waveform, but also cause the current amplitude and phase. There are obvious changes in the distribution law, so this view feature plays a key supporting role in fault identification and analysis [8,10].

The auxiliary Z-score view provides another dimensional characterization perspective by highlighting the local waveform and relative timing changes. The view weakens the interference caused by the difference in absolute amplitude scale, so that the model can focus on the law of waveform evolution, transient fluctuation characteristics, and local pattern changes. From this, the model can synchronously learn the amplitude characteristics related to the fault intensity and the structural characteristics related to the waveform morphology from the same set of three-phase current signals. The results of the ablation experiment show that the improvement of diagnostic performance is not simply due to the increase in the number of input channels, but rather to the introduction of a dual-view characterization method with both physical connotation and statistical complementarity. This also explains the mechanism. Compared with the variant model that only uses physical-range normalization or basic residual architecture, the proposed method has better overall fault classification performance.

The feature visualization results further support this interpretation. Compared with single-view or baseline representations, the dual-view representation tends to produce more separable feature distributions among different operating conditions. This is beneficial for distinguishing composite grounding faults and normal states, which usually show relatively clear differences in current magnitude and phase relationship. Meanwhile, the remaining misclassifications between B-G, C-G, and adjacent classes imply that some weak or boundary fault cases may still share similar local patterns, especially under limited-sample conditions. Therefore, the dual-view representation improves the discriminative capacity of the model, but further enhancement is still needed for fault cases with subtle inter-class differences.

4.2. Influence of the Robust Training Strategy

The results also show that a robust training strategy helps to achieve stable model optimization under small sample conditions. In actual cable monitoring, it is usually difficult to obtain a large number of labeled fault samples, especially for composite faults, weak grounding abnormalities, and early defects. In this case, deep neural networks may be easier to fit the fluctuations of specific samples than to learn generalizable failure patterns. The training strategy adopted in this study is intended to alleviate this issue by combining Mixup, label smoothing, dropout, and cosine annealing, so that the model can learn smoother decision-making boundaries from limited training data.

Considering the limited number of simulated samples and the continuous variation in sheath loop current characteristics under different fault conditions, Mixup augmentation is adopted in this study. The method constructs additional training samples by linearly combining original samples and their corresponding labels. This process enables the model to learn transitional states between different operating conditions and reduces the risk of overfitting to individual simulated samples [35]. To further improve training stability, label smoothing is introduced to soften the target label distribution and prevent the classifier from becoming overly confident in a single class [37]. Dropout is then used to weaken

the excessive dependence among high-level features, while cosine annealing is applied to adjust the learning rate in a smooth manner during the training process [38,39].

The ablation results indicate that the improvement of the proposed method cannot be explained by any single component. Adding physical-range normalization, Mixup, or cosine annealing alone brings only a limited increase compared with the basic residual model. The highest performance is obtained when the dual-view input representation and the robust training strategy are used together. This suggests that the proposed method should be viewed as a complete diagnostic framework, rather than a direct combination of several existing deep-learning techniques. In this framework, the input design, network structure, and training strategy support each other, which helps the model maintain better stability and generalization when only limited data are available.

4.3. Engineering Implications for Cable Online Monitoring

In renewable-powered distribution systems and microgrids, cable feeders and interconnection lines are critical assets for maintaining operational continuity; therefore, component-level monitoring of sheath current abnormalities can provide useful support for resilient operation and maintenance decision-making. From an engineering point of view, the proposed method offers a practical route for intelligent monitoring of cable sheath circulating current signals. It takes three-phase time-series current signals as the direct input, without converting them into two-dimensional time-frequency images. This simplifies the preprocessing procedure and helps retain the original temporal information in the current signals. In addition, the lightweight 1D-ResNet structure is more suitable for online monitoring than large deep networks with high computational demand. This is important for cable monitoring systems, where diagnostic models need to respond quickly under limited computing resources.

The six-channel dual-view input also improves the interpretability of the diagnostic process. The physically normalized channels mainly reflect magnitude-related information, while the z-score channels focus more on waveform morphology. Compared with purely data-driven feature construction, this design is closer to the physical behavior of sheath circulating current faults. In engineering applications, it may help explain why the model can distinguish different fault types, especially when abnormal sheath current is related to changes in grounding state, phase coupling, or current distribution.

The proposed method is also consistent with the development of condition-based maintenance and early warning for power cable systems. After being integrated with online sheath current sensors, the model could be used to identify abnormal operating states and provide support for maintenance decisions before faults develop into serious insulation degradation or line outages. However, further validation is still needed before practical deployment. The model should be tested under different cable lengths, grounding modes, load levels, noise conditions, and sensor errors. Real-time applications also require model compression, inference-time analysis, and threshold-based warning rules, so that the diagnostic results can be effectively used in maintenance workflows.

4.4. Limitations and Future Work

The proposed method still has several limitations. The current validation is mainly based on an ATP-EMTP-generated dataset. This dataset is controllable and repeatable, which is useful for comparing diagnostic methods when real cable fault samples are difficult to obtain. Simulation data cannot fully reflect the operating conditions of actual cable systems. In practical applications, sheath circulating current signals may be affected by load changes, grounding resistance variation, electromagnetic interference, sensor drift,

installation differences, and environmental factors. The proposed method still needs to be tested using measured data from laboratory platforms and field monitoring systems.

The dataset used in this study contains only 40 labeled samples, with five samples for each operating condition. Although a unified 5×5 repeated stratified cross-validation protocol was used to reduce the influence of data partitioning, the reported performance may still be affected by the limited number of samples in each class. Therefore, the current results should be interpreted as preliminary feasibility evidence under a small-sample simulation setting rather than a definitive estimate of field performance. In future work, more operating conditions should be added, such as different grounding resistances, fault inception angles, load levels, cable layouts, and noise levels. Independent test sets and larger measured datasets are also needed to further examine model stability.

The confusion matrix shows that some B-G and C-G single-phase grounding faults are still misclassified. Some weak fault cases may have current features close to those of nearby classes. The model, therefore, still has difficulty distinguishing fault states with similar current distributions or small waveform differences. Later studies may consider attention mechanisms, metric learning, or physics-based constraints to improve the separation of similar fault classes. The current robustness test only considers 12% relative Gaussian noise. Although the proposed method maintained competitive accuracy under the tested 12% relative Gaussian noise condition, the current robustness test only considers one Gaussian-noise setting. It does not fully represent structured disturbances in renewable-powered microgrid environments, such as non-Gaussian impulsive noise, harmonic distortion, inverter-driven switching noise, sensor bias, missing data, and load fluctuation. These disturbance types should be further considered so that the test conditions are closer to practical online monitoring environments.

Further work should focus on measured-data validation, dataset expansion, and engineering deployment. The proposed framework needs to be tested using measured sheath circulating current data from laboratory platforms or field monitoring systems. More fault scenarios and operating disturbances should be included so that the model can be evaluated under conditions closer to practical cable operation. For online use, computational cost, inference speed, and warning threshold design also need further study.

5. Conclusions

This paper proposes a dual-view Mixup-ResNet method for fault diagnosis of cable sheath circulating current signals, with a particular focus on intelligent monitoring of critical cable assets in renewable-driven power systems. By integrating a physically normalized primary view with a sample-wise auxiliary z-score view, the proposed approach constructs a six-channel input representation that simultaneously preserves fault-intensity information and waveform-shape characteristics. On this basis, a lightweight 1D-ResNet is combined with Mixup augmentation and robust training strategies, including label smoothing, dropout, and cosine annealing, to achieve reliable fault recognition under small-sample conditions.

Experimental results on the ATP-EMTP-based eight-class dataset demonstrate the preliminary effectiveness of the proposed method. Under the unified 5×5 repeated stratified cross-validation protocol, the proposed model achieves an average diagnostic accuracy of 91.50%, a weighted F1-score of 91.69%, and a macro F1-score of 91.69%. These results provide a more conservative and internally consistent estimate of the proposed method under the current small-sample simulation setting. The ablation results further indicate that the performance improvement is mainly derived from the combined contribution of the dual-view input representation, residual feature extraction, and Mixup-based regularization strategy.

Under 12% relative Gaussian noise, the proposed method achieves an average accuracy of 91.0%. This result suggests that the model retains competitive noise tolerance under the tested Gaussian-noise condition, but it should not be interpreted as evidence of superior robustness under all practical disturbances. More comprehensive robustness evaluation under field-like noise, load fluctuation, sensor bias, and other operating disturbances are still required.

Overall, the present work should be regarded as a simulation-based feasibility study for small-sample fault diagnosis of cable sheath circulating current signals. Although the results suggest the potential of the proposed framework, the current validation is still limited to ATP-EMTP-generated data and a limited sample scale. Further validation using laboratory measurements and field monitoring data is required before practical deployment. Future work will focus on measured-data validation, dataset expansion, and evaluation under more realistic operating disturbances.

In addition, transfer learning, semi-supervised learning, and unsupervised domain adaptation will be explored to reduce the dependence on large-scale labeled fault samples [43,44]. These strategies are particularly important for practical XLPE cable sheath monitoring because real sheath grounding faults are rare and difficult to label in large quantities under field operating conditions. By transferring knowledge from simulation data, laboratory data, or related cable operating conditions to field monitoring scenarios, the proposed framework may be further adapted to practical applications with limited labeled fault data.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/en19143290/s1>. Supplementary File S1: Source code package for the proposed dual-view Mixup-ResNet method, including the README file, dependency list, input-format example, class-naming rules, random-seed settings, and execution instructions.

Author Contributions: Conceptualization, H.Y. and X.L.; methodology, H.Y.; software, H.Y. and J.H.; validation, H.Y., J.H. and B.Z.; formal analysis, H.Y. and B.Z.; investigation, H.Y. and B.Z.; resources, X.L. and J.M.; data curation, H.Y.; writing—original draft preparation, H.Y.; writing—review and editing, X.L., J.M. and B.Z.; visualization, H.Y.; supervision, X.L.; project administration, X.L.; funding acquisition, X.L. and H.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Key R&D Program of Zhejiang, grant number 2025C02001, and the Zhejiang Provincial College Student Science and Technology Innovation Activity Plan (Xinmiao Talent Program), grant number 2026R409B46.

Data Availability Statement: The source code for the algorithms and computational workflow used in this study is provided in Supplementary File S1, together with a README file, dependency list, input-format example, class-naming rules, random-seed settings, and execution instructions. The processed current records and original ATP-EMTP/ATPDraw project files are not publicly released because of institutional data-management restrictions and software-version dependency. The principal simulation and sample-generation settings are reported in Tables 1 and 2. The processed data may be made available by the corresponding author upon reasonable request.

Conflicts of Interest: Author Jinwei MAO was employed by Shaoxing Shangyu Hangxie Thermal Power Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

XLPE	Cross-Linked Polyethylene
ATP-EMTP	Alternative Transients Program–Electromagnetic Transients Program
CNN	Convolutional Neural Network
LSTM	Long Short-Term Memory
ResNet	Residual Network
SVM	Support Vector Machine
KNN	K-Nearest Neighbors

References

1. Mahmoud, M.A.; Nasir, N.R.M.; Gurunathan, M.; Raj, P.; Mostafa, S.A. The Current State of the Art in Research on Predictive Maintenance in Smart Grid Distribution Network: Fault's Types, Causes, and Prediction Methods-A Systematic Review. *Energies* **2021**, *14*, 5078. [\[CrossRef\]](#)
2. Alhanaf, A.S.; Balik, H.H.; Farsadi, M. Intelligent Fault Detection and Classification Schemes for Smart Grids Based on Deep Neural Networks. *Energies* **2023**, *16*, 7680. [\[CrossRef\]](#)
3. Salehimehr, S.; Miraftabzadeh, S.M.; Brenna, M. A Novel Machine Learning-Based Approach for Fault Detection and Location in Low-Voltage DC Microgrids. *Sustainability* **2024**, *16*, 2821. [\[CrossRef\]](#)
4. Ghaforian, A.; Duggan, P.; Lu, L. A Comprehensive Review of Cable Monitoring Techniques for Nuclear Power Plants. *Energies* **2025**, *18*, 2333. [\[CrossRef\]](#)
5. Tao, J.; Rehman, S.U.; Ali, R.; Raza, S.A. Advancement and Challenges: A Review of Power Cable Aging Monitoring and Diagnostic Techniques. *Renew. Sustain. Energy Rev.* **2025**, *222*, 115970. [\[CrossRef\]](#)
6. Bindi, M.; Piccirilli, M.C.; Luchetta, A.; Grasso, F. A Comprehensive Review of Fault Diagnosis and Prognosis Techniques in High Voltage and Medium Voltage Electrical Power Lines. *Energies* **2023**, *16*, 7317. [\[CrossRef\]](#)
7. Lu, B.; Li, S.; Cui, Y.; Zhao, X.; Zhang, D.; Kang, Y.; Dong, H. Insulation Degradation Mechanism and Diagnosis Methods of Offshore Wind Power Cables: An Overview. *Energies* **2023**, *16*, 322. [\[CrossRef\]](#)
8. Li, G.; Chen, J.; Li, H.; Hu, L.; Zhou, W.; Zhou, C. Detection of Irregular Sheath Current Distribution for Diagnosis of Faults in Grounding Systems of Cross-Bonded Cables. *IEEE Access* **2023**, *11*, 68453–68461. [\[CrossRef\]](#)
9. Li, G.; Zhou, W.; Zhou, C.; Yi, J.; Qiu, L.; Ai, Y. Fault Diagnosis of Grounding System of High Voltage Cable Circuits Using Graph Attention Networks. *Int. J. Electr. Power Energy Syst.* **2025**, *168*, 110683. [\[CrossRef\]](#)
10. Wan, Q.; Yan, X. Fault Diagnosis of HV Cable Metal Sheath Grounding System Based on LSTM. *Appl. Sci.* **2023**, *13*, 2453. [\[CrossRef\]](#)
11. Zhang, Q.; Li, K.; Chen, L.; Luo, J.; Zhao, Z. Identification and Localization Study of Grounding System Defects in Cross-Bonded Cables. *Electronics* **2025**, *14*, 622. [\[CrossRef\]](#)
12. Ai, Y.; Song, B.; Wu, S.; Li, Y.; Lu, L.; Wang, L. Diagnosis of Reverse-Connection Defects in High-Voltage Cable Cross-Bonded Grounding System Based on ARO-SVM. *Sensors* **2025**, *25*, 590. [\[CrossRef\]](#) [\[PubMed\]](#)
13. Yuan, W.; Tu, J.; Ai, Y.; Xia, Z.; Song, R.; He, J.; Gao, X.; Jiang, M.; Wang, H. Two-Sheath Loop Short Circuit Defects Detection in High-Voltage Cable Systems Using Sheath Current Phasors. *Energies* **2025**, *18*, 4868. [\[CrossRef\]](#)
14. Yang, H.; Wang, J.; Zhao, P.; Yu, C.; You, H.; Dou, J. On-Line Insulation Monitoring Method of Substation Power Cable Based on Distributed Current Principal Component Analysis. *Energies* **2025**, *18*, 688. [\[CrossRef\]](#)
15. Huo, Y.; Wang, K.; Lampe, L.; Leung, V.C.M. Validation of Machine Learning-Aided and Power Line Communication-Based Cable Monitoring Using Measurement Data. *Sensors* **2024**, *24*, 335. [\[CrossRef\]](#) [\[PubMed\]](#)
16. Shao, Q.; Fan, S.; Zhang, Z.; Liu, F.; Fu, Z.; Lv, P.; Mu, Z. Artificial Intelligence in Cable Fault Detection and Localization: Recent Advances and Research Challenges. *Energies* **2025**, *18*, 3662. [\[CrossRef\]](#)
17. Dong, X.; Yang, Y.; Zhou, C.; Hepburn, D.M. Online Monitoring and Diagnosis of HV Cable Faults by Sheath System Currents. *IEEE Trans. Power Deliv.* **2017**, *32*, 2281–2290. [\[CrossRef\]](#)
18. Marzinotto, M.; Mazzanti, G. The Feasibility of Cable Sheath Fault Detection by Monitoring Sheath-to-Ground Currents at the Ends of Cross-Bonding Sections. *IEEE Trans. Ind. Appl.* **2015**, *51*, 5376–5384. [\[CrossRef\]](#)
19. Hernández-Mayoral, E.; Madrigal-Martínez, M.; Mina-Antonio, J.D.; Iracheta-Cortez, R.; Enríquez-Santiago, J.A.; Rodríguez-Rivera, O.; Martínez-Reyes, G.; Mendoza-Santos, E. A Comprehensive Review on Power-Quality Issues, Optimization Techniques, and Control Strategies of Microgrid Based on Renewable Energy Sources. *Sustainability* **2023**, *15*, 9847. [\[CrossRef\]](#)
20. Abbas, A.S.; El-Sehiemy, R.A.; Abou El-Ela, A.A.; Ali, E.S.; Mahmoud, K.; Lehtonen, M.; Darwish, M.M.F. Optimal Harmonic Mitigation in Distribution Systems with Inverter Based Distributed Generation. *Appl. Sci.* **2021**, *11*, 774. [\[CrossRef\]](#)

21. Zhang, W.; Song, Y.; Wu, X.; Liu, H.; Tian, H.; Tang, Z.; Xu, S.; Chen, W. Detecting Partial Discharge in Cable Joints Based on Implanting Optical Fiber Using MZ-Sagnac Interferometry. *Sensors* **2025**, *25*, 3166. [[CrossRef](#)] [[PubMed](#)]
22. Navarrete-Rajadel, P.; Llovera-Segovia, P.; Fuster-Roig, V.; Quijano-López, A. A Sensor for Multi-Point Temperature Monitoring in Underground Power Cables. *Sensors* **2025**, *25*, 5490. [[CrossRef](#)] [[PubMed](#)]
23. Pan, W.; Chen, X.; Zhao, K. Cable-Partial-Discharge Recognition Based on a Data-Driven Approach with Optical-Fiber Vibration-Monitoring Signals. *Energies* **2022**, *15*, 5686. [[CrossRef](#)]
24. Zhu, W.; Hui, B.; Li, J.; Han, T.; Zhao, L.; Hou, S. Frequency Domain Reflectometry for Power Cable Defect Localization: A Comparative Study of FFT and IFFT Methods. *Energies* **2025**, *18*, 5346. [[CrossRef](#)]
25. Wang, R.; Pan, R. A Cable Defect Assessment Method Based on a Mixed-Domain Multi-Feature Set of Overall Harmonic Signals. *Energies* **2025**, *18*, 83. [[CrossRef](#)]
26. Fu, H.; Ai, Y.; Qiu, L.; Ai, Y.; Yang, B.; Shi, Y. Application of Fuzzy Comprehensive Evaluation Method in the Identification of Potential Faults of High-Voltage Power Cables. *Front. Energy Res.* **2024**, *12*, 1490524. [[CrossRef](#)]
27. Shadi, M.R.; Mirshekali, H.; Shaker, H.R. Partial Discharge-Based Cable Vulnerability Ranking with Fuzzy and FAHP Models: Application in a Danish Distribution Network. *Sensors* **2025**, *25*, 3454. [[CrossRef](#)] [[PubMed](#)]
28. Wang, F.; Zhang, P.; Li, J.; Li, Z.; Zhao, M.; Liang, Y.; Su, G.; You, X. Multi-Feature Based Extreme Learning Machine Identification Model of Incipient Cable Faults. *Front. Energy Res.* **2024**, *12*, 1364528. [[CrossRef](#)]
29. Kumar, H.; Shafiq, M.; Kauhaniemi, K.; Elmusrati, M. A Review on the Classification of Partial Discharges in Medium-Voltage Cables: Detection, Feature Extraction, Artificial Intelligence-Based Classification, and Optimization Techniques. *Energies* **2024**, *17*, 1142. [[CrossRef](#)]
30. Li, Y.; Gong, C.; Deng, T.; Jia, Z.; Wang, F.; Zhao, Q.; Zhang, J. Partial Discharge Pattern Recognition Based on Swin Transformer for Power Cable Fault Diagnosis in Modern Distribution Systems. *Processes* **2025**, *13*, 852. [[CrossRef](#)]
31. Zhang, L.-X.; Zhou, Y.-Y.; Yao, S.-J.; Chai, J.-L.; Chen, Y.-J.; Zhang, Z.-S. Real Defect Partial Discharge Identification Method for Power Cables Joints Based on Integrated PJS-M and GA-SVM Algorithm with Multi-Source Fusion. *Front. Energy Res.* **2025**, *13*, 1622318. [[CrossRef](#)]
32. Han, W.; Yang, G.; Hao, C.; Wang, Z.; Kong, D.; Dong, Y. A Data-Driven Model of Cable Insulation Defect Based on Convolutional Neural Networks. *Appl. Sci.* **2022**, *12*, 8374. [[CrossRef](#)]
33. Moutassem, Z.; Bounaim, D.; Li, G. A Physics-Enhanced CNN-LSTM Predictive Condition Monitoring Method for Underground Power Cable Infrastructure. *Algorithms* **2025**, *18*, 600. [[CrossRef](#)]
34. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778. [[CrossRef](#)]
35. Zhang, H.; Cisse, M.; Dauphin, Y.N.; Lopez-Paz, D. Mixup: Beyond Empirical Risk Minimization. In Proceedings of the 6th International Conference on Learning Representations (ICLR), Vancouver, BC, Canada, 30 April–3 May 2018. [[CrossRef](#)]
36. Zhao, Y.; Liu, Q.; Xia, R.; Shang, Y. Active Cable Sheath Current-Based Protection Scheme for an Offshore HVAC Wind Transmission System. *Front. Energy Res.* **2022**, *10*, 844474. [[CrossRef](#)]
37. Szegedy, C.; Vanhoucke, V.; Ioffe, S.; Shlens, J.; Wojna, Z. Rethinking the Inception Architecture for Computer Vision. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 2818–2826. [[CrossRef](#)]
38. Loshchilov, I.; Hutter, F. SGDR: Stochastic Gradient Descent with Warm Restarts. In Proceedings of the 5th International Conference on Learning Representations (ICLR), Toulon, France, 24–26 April 2017.
39. Srivastava, N.; Hinton, G.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *J. Mach. Learn. Res.* **2014**, *15*, 1929–1958.
40. Kingma, D.P.; Ba, J. Adam: A Method for Stochastic Optimization. In Proceedings of the 3rd International Conference on Learning Representations (ICLR), San Diego, CA, USA, 7–9 May 2015. [[CrossRef](#)]
41. Wu, K.; Hao, S.; Liang, Y.; Yang, B.; Huang, Z. Target Recognition for Ultra-Wideband Radio Fuzes Using 1D-CGAN-Augmented 1D-CNN. *Entropy* **2025**, *27*, 980. [[CrossRef](#)] [[PubMed](#)]
42. Tchatchoua, P.; Graton, G.; Ouladsine, M.; Christaud, J.-F. Application of 1D ResNet for Multivariate Fault Detection on Semiconductor Manufacturing Equipment. *Sensors* **2023**, *23*, 9099. [[CrossRef](#)] [[PubMed](#)]
43. Zhao, Z.; Zhang, Q.; Yu, X.; Sun, C.; Wang, S.; Yan, R.; Chen, X. Applications of Unsupervised Deep Transfer Learning to Intelligent Fault Diagnosis: A Survey and Comparative Study. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 3525828. [[CrossRef](#)]
44. Wang, Q.; Taal, C.; Fink, O. Integrating Expert Knowledge With Domain Adaptation for Unsupervised Fault Diagnosis. *IEEE Trans. Instrum. Meas.* **2022**, *71*, 3500312. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.