# An Explainable and Robust Method for Fault Classification and Location on Transmission Lines

Jiashu Fang, Kunjin Chen, Chongru Liu, *Senior Member, IEEE*, and Jinliang He, *Fellow, IEEE*

***Abstract*—Machine learning-based approaches for fault diagnosis on transmission lines have attracted increasing attention in recent years, yet concerns over their robustness and explainability hamper their practical applications. Moreover, separate algorithms are normally used to solve the fault classification and location tasks. In this work, we focus on these tasks and propose an integrated model based on the convolutional neural network, whose improved performance over separate models is demonstrated by numerical results. Besides, explainability of the proposed model is demonstrated and analyzed. Specifically, the class activation maps and attention maps illustrate that the proposed structure can reduce the degradation of the model performance due to data pollution, enhancing the credibility of the proposed method in practical applications. Moreover, the proposed model outperforms existing machine learning-based models in terms of accuracy and robustness. Such a structure has been proven effective on field data and can be applied to many other tasks of fault diagnosis.**

***Index Terms*—Convolutional networks, explainability, fault classification, fault location, robustness, transmission lines.**

## I. Introduction

WITH the accelerated growth of power loads and the large-scale integration of renewable energies, long-distance, and bulk-capacity power transmission has become the trend of future power grids [1]. However, the increased voltage level of transmission lines exacerbates the negative impact of outages caused by transmission line faults, which account for more than 85% of power system faults [2]. As a result, rapid and accurate fault classification and location on transmission lines are critical for power system reliability [3]. Although circuit breakers and relays are widely deployed to monitor the status of transmission lines, their mis-operation occasionally occurs, causing serious cascading failures or outages [4]. In addition, reliance on the knowledge of transition resistance [5] and the accurate recognition of initial wave head [6], respectively, limits the traditional impedance-based and traveling-wave-based methods for fault location estimation. With the collaborative operation of wireless sensor networks (WSNs) and the development of edge computing units in smart grids [7], the reliable and real-time information on transmission lines can be monitored and further analyzed, motivating the development of specialized data-driven fault-diagnosis methods for fault classification and location applications.

As for data-driven fault classification and location tasks, there exist two major categories of methods in the literature. The first category generally has two steps, namely, feature extraction and classification. To extract features from faulty voltage and current signals, various frequency-domain techniques, such as discrete Fourier transform [8] and discrete wavelet transform [9], [10], have been used. The subsequent classification methods are developed using different data-driven approaches, including multilayer perceptron (MLP) [10], extreme learning machines [11], fuzzy-logic systems [12], support vector machines [9], [13], and random forests (RF) [8], [14]. An ensemble learning algorithm is developed in [14], where the RF algorithm is trained with the incremental quantities of fault current signals that are calculated using moving average filters. However, the two-step methods highly rely on expert knowledge when extracting the features and lack closed-loop feedback between feature extraction and model results.

The other category addresses the fault diagnosis in an end-to-end structure, which reduces the reliance on expert knowledge and commonly leads to the better performance owing to the iterative optimization of model parameters. These deep learning methods attempt to find fault signal representations that contain more useful information in a data-driven manner. In [15], denoising sparse autoencoders (DSAEs) are used to extract features from data through a reconstruction task. These features are then used as filters in the convolution operation [16] with signals, producing representatives for the signals. The group sparse

Manuscript received 20 October 2022; accepted 29 November 2022. Date of publication 30 January 2023; date of current version 11 August 2023. This work was supported by the National Natural Science Foundation of China under Grant 51921005. Paper no. TII-22-4365. *(Corresponding authors: Chongru Liu; Jinliang He.)*

Jiashu Fang and Chongru Liu are with the State Key Laboratory of Alternate Electrical Power System With Renewable Energy Sources, North China Electric Power University, Beijing 102206, China (e-mail: jiashu.fang@hotmail.com; chongru.liu@ncepu.edu.cn).

Kunjin Chen is with the Alibaba Group, Hangzhou 311121, China (e-mail: yalickj@163.com).

Jinliang He is with the State Key Lab of Power Systems and Department of Electrical Engineering, Tsinghua University, Beijing 100084, China (e-mail: hejl@tsinghua.edu.cn).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TII.2022.3229497.

Digital Object Identifier 10.1109/TII.2022.3229497

representation classification (GSRC) is proposed in [17], where the fault type is determined by comparing the reconstruction error of different fault-type-specific groups in the training samples dictionary. Qin et al. [18] proposed an online fault diagnosis method based on the deep belief neural network, which extracts and classifies the characteristics of fault signals automatically. Through image processing, Lei and Sui [19] achieved highly accurate fault classification of high-voltage lines based on the region convolutional neural network (CNN). In [20], an adaptive CNN-based method is proposed for fault location, and the pooling model is improved to increase its feature extraction ability. Belagoune et al. [21] extracted spatiotemporal sequences of high-dimensional electrical features by building two separate models using the long short-term memory (LSTM) structure, which presents high accuracy in both fault classification and location tasks.

Nonetheless, the above data-driven methods exhibit limitations when applied to practical power systems. First, the online data collected by WSNs contain various noises [22], yet the previous methods rarely consider robustness against data pollution when designing the structure of the algorithm, hampering the application of data-driven models to real-world scenarios. Second, explainability is lacking in existing algorithms, leading to concerns from grid operators about the fault diagnosis decisions given by the artificial intelligence algorithms, especially the end-to-end deep learning method. Third, the previous approaches (e.g., [9], [21], [23], [24]) commonly utilize separate algorithms for fault classification and location based on the post-fault transient voltage and/or current waveforms, which are jointly decided by the fault type and fault distance in transmission lines. Since the high relevance between the fault classification and location tasks is not efficiently utilized by the separate models, the performance and robustness of data-driven models are limited [25].

Considering the existing research gaps, in this work, we proposed an explainable and robust integrated model for fault classification and location. As we use the postfault transient voltage or current waveforms as the model's input for fault diagnosis, the faulty temporal series data would be stored as high-dimensional arrays, which present data structures similar to images. Accordingly, this work adopts the CNN, which is a widely used model for image classification. Compared with FC neural networks [24], the utilization of CNN greatly reduces the number of model parameters as the convolutional neurons share the same weights for each spatial location when extracting the high-dimensional data [26]. In order to improve model robustness against data pollution, the nonlocal block [27] and DropBlock [28] are utilized to, respectively, capture the long-range dependencies and temporally dispersed fault information in samples. Specifically, the attention operation in nonlocal blocks introduces effective multihop dependency modeling ability into the proposed method to automatically neglect consecutive measurement errors [27], which are widely witnessed in fault signals due to stochastic packet drops and communication failures in data collection units [22]. The main contributions of this article include the following.

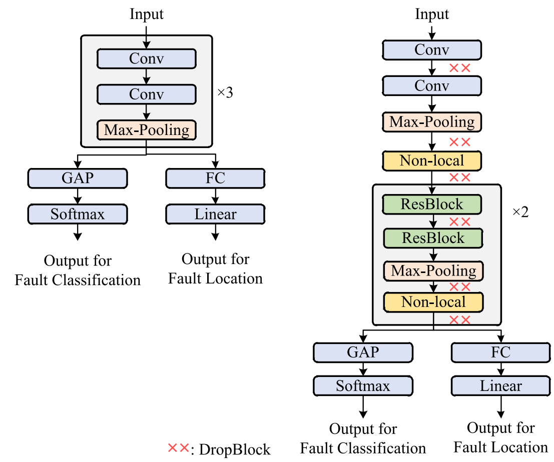



Fig. 1. Fault diagnosis models. Left: the benchmark model with three convolutional sections. Right: the proposed model.

1) The model structure is designed to improve robustness against data pollution, which has been verified with different levels of data pollution and field data, demonstrating the feasibility of the proposed model in real applications.
2) Model explainability is realized based on the proposed structure for fault diagnosis. This not only visualizes and proves the model robustness, but also enhances the credibility of the data-driven method in practice.
3) A multitask framework is developed to introduce the relevance between fault classification and location tasks as an inductive bias, leading to better robustness than separate models.

The rest of this article is organized as follows. The proposed convolutional network model is elaborated in Section II. In Section III, we will analyze the proposed model in terms of robustness and explainability and compare its performance versus existing approaches. Verification of the proposed model on field data is also provided. Finally, Section IV concludes this article.

## II. Proposed Fault Classification and Location Approach

In this section, we propose a multitask model that outputs fault classification and location estimation via a single framework. Several components are utilized to modify the benchmark model, effectively increasing the model's robustness and explainability.

### A. Proposed Model Overview

The structure of the proposed network is shown in Fig. 1 (right). Several components, including residual blocks (Resblocks) [29], nonlocal blocks, and DropBlock layers, are added to the benchmark model (a convolutional network) shown in Fig. 1 (left). A similar model structure has been used for fault diagnosis on transmission lines [30].

For both models, global average pooling (GAP) and fully connected (FC) layers are added at the end of the structure, and the outputs for fault classification and location are separately estimated through softmax [31] and linear functions. The application of GAP layers enables the visualization of class activation maps (CAMs) for explainability. The fault location under the nonfaulty scenario is set to 0 during the training process, and gives a nonzero output once the fault classification determines an abnormal operation during the test process or actual application. The loss weights in the fault location and classification tasks are set to 0.8 and 0.2. A typical feature map at the $l$th convolutional layer of the network has the form $\mathbf{x}^l \in \mathbb{R}^{T^l \times C^l}$, where $T^l$ is the number of timesteps and $C^l$ is the number of channels of the $l$th convolutional layer. In this work, $T^1 = 200$ and $C^1 = 6$ since the input sample has 200 timesteps and 6 signals, as will be discussed in Section III-A. For simplicity, the superscripts for layer indexes are omitted hereinafter. A detailed explanation of the convolutional and pooling layers for 1-D multichannel signals can be found in [15].

To simplify the optimization process in deep networks, Resblocks [29] are added to the proposed structure. A typical residual block has the form

$$\mathbf{y} = f_{\text{res}}(\mathbf{x}) + \mathbf{x} \tag{1}$$

where $f_{\text{res}}$ is the transformation function in the residual block. In this work, we simply use a convolutional layer for $f_{\text{res}}$. The "$+\mathbf{x}$" part in (1) is the residual connection. When the input and output feature maps have different numbers of channels, a convolutional layer is also added to this path so that the two terms in the equation add in an element-wise manner.

## B. Components for Robustness

In this work, the nonlocal block and the DropBlock are added to the model for better robustness. Specifically, the nonlocal block enables the model to integrate fault-related information within the feature maps (especially for information within timesteps that are far from each other), and the DropBlock layer forces the model to collect more fault-related information that is temporally dispersed.

The nonlocal block allows the features at each time position in a feature map to interact with features at all timesteps rather than timesteps in a local region [27]. With reference to the diagram of the nonlocal block in Fig. 2, the nonlocal operation for an input feature map $\mathbf{x} \in \mathbb{R}^{T \times C_{\text{in}}}$ writes

$$\mathbf{y}_i = \frac{1}{T}\sum_{j=1}^{T} f(\mathbf{x}_i, \mathbf{x}_j) g(\mathbf{x}_j) \tag{2}$$

where $T$ is the timestep number of the feature map, $\mathbf{y}$ is the output feature map, and the subscripts $i$ and $j$ of $\mathbf{x}$ and $\mathbf{y}$ are indexes of timesteps. The pairwise function $f$, which takes feature maps of two timesteps as input and produces a scalar, is defined as

$$f(\mathbf{x}_i, \mathbf{x}_j) = \theta(\mathbf{x}_i)^{\mathrm{T}} \phi(\mathbf{x}_j) \tag{3}$$

where $\theta(\mathbf{x}_i) = \mathbf{W}_\theta \mathbf{x}_i$, $\phi(\mathbf{x}_j) = \mathbf{W}_\phi \mathbf{x}_j$, and $\mathbf{W}_\theta$ and $\mathbf{W}_\phi$ are weight matrices to be learned. Similarly, $g(\mathbf{x}_j) = \mathbf{W}_g \mathbf{x}_j$ and $\mathbf{W}_g$ is a weight matrix. In Fig. 2, $\theta(\mathbf{x}_i)$, $\phi(\mathbf{x}_j)$, and $g(\mathbf{x}_j)$ are named as the *query* of timestep $i$, the *key* of timestep $j$, and the *value* of timestep $j$, respectively. For a given timestep $i$, the inner product of its query $\theta(\mathbf{x}_i)$ and the key of each timestep is calculated to produce the attention scores for all timesteps. Specifically, normalizing $f(\mathbf{x}_i, \mathbf{x}_j)$ with $\frac{1}{T}$ gives the attention score assigned to $g(\mathbf{x}_j)$ for timestep $i$. As shown in Fig. 2, $\theta$, $\phi$, and $g$ are implemented as convolution operations of $C'$ filters with a filter size of one along the timesteps.

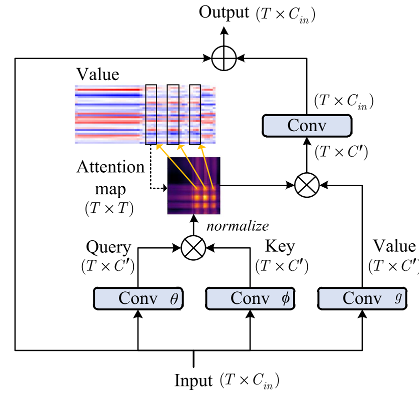



Fig. 2. Nonlocal block for 1-d feature maps. An attention map is obtained by mapping the input feature map into two embedding spaces, and by using the dot-product similarity. "⊗" and "⊕" stand for matrix multiplication and element-wise sum, respectively.

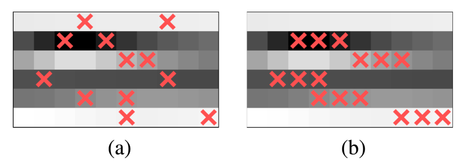



Fig. 3. Two regularization methods based on randomly dropping values in feature maps. (a) Dropout. (b) DropBlock. For 1-d feature maps with multiple channels, DropBlock is performed within each channel.

The overall nonlocal block is then further defined as

$$\mathbf{z}_i = \mathbf{W}_z \mathbf{y}_i + \mathbf{x}_i \tag{4}$$

where $\mathbf{z}$ is the output of the block and $\mathbf{W}_z$ transforms $\mathbf{y}_i$ so that it matches the size of the input (i.e., from $C'$ to $C_{\text{in}}$). The residual connection from $\mathbf{x}_i$ ensures that local information is also found in the output of the block.

DropBlock is another component that can increase robustness of the model [28]. An illustration of DropBlock is shown in Fig. 3. Unlike the dropout method that randomly samples a mask $M : M_{i,j} \sim \text{Bernoulli}(p)$ and drops the values (i.e., set the values to 0) where $M_{i,j} = 0$, DropBlock further sets the values surrounding the 0 positions in $M$. In the 1-D case, the blocks being dropped are restricted within channels. Specifically, after $p_{\text{keep}}$ (the probability a unit in the feature map is kept) and $S_{\text{b}}$ (the size of the blocks) are determined, the parameter $\gamma$ is calculated as

$$\gamma = \frac{1 - p_{\text{keep}}}{S_{\text{b}}} \times \frac{T}{T - S_{\text{b}} + 1}. \tag{5}$$

The center points of the blocks to be dropped are the positions of zeros in $M : M_{i,j} \sim \text{Bernoulli}(\gamma)$. During training time, the masks are sampled for each sample, and the feature maps are multiplied by $N_{\text{feat}}/(N_{\text{feat}} - N_{\text{drop}})$, where $N_{\text{feat}}$ is the number

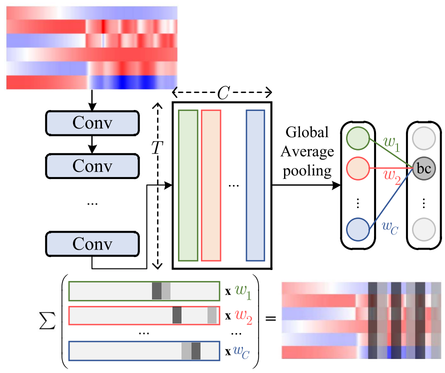



Fig. 4. Diagram explaining how the CAMs are generated. The CAM reveals the most important regions for fault diagnosis.

of units in the feature map, and $N_{\text{drop}}$ is the number of units dropped. As fault-related transients appear in temporally contiguous sections within a signal, DropBlock is more likely to mask out clues that can be used to infer the fault classification and location. Thus, the model is forced to collect more information from temporally distributed areas based on the input and the feature maps. Such a regularization increases robustness of the model.

## C. Components for Explainability

Explainability of the model is obtained by assessing the contribution of different input features to the output results, and helps identify the most important regions in the database for specific tasks. In this article, the modified explainable component is also used to verify the model robustness.

Explainability of the proposed convolutional network model can be achieved by CAMs [32]. The procedure for generating CAMs is illustrated in Fig. 4. Specifically, the activation of the $c$th GAP unit is calculated as $a_c = \frac{1}{T}\sum_{i=1}^{T} x_i^c$, which simply averages the values of all units for the channel, where $x_i^c$ is the activation of the unit at the $i$th timestep and the $c$th channel in the feature map before the GAP layer.

The input to the softmax function at the output layer for class $k$ then becomes

$$q_k = \sum_{c=1}^{C} w_c^k a_c = \sum_{c=1}^{C} w_c^k \cdot \frac{1}{T}\sum_{i=1}^{T} x_i^c = \frac{1}{T}\sum_{i=1}^{T}\sum_{c=1}^{C} w_c^k x_i^c \quad (6)$$

where $w_c^k$ is the weight associating the $c$th GAP unit and the $k$th class at the output. Finally, the output of the softmax function for class $k$ is $p_k = \exp(q_k)/\sum_l \exp(q_l)$. In order to obtain the CAM for a given input, we define the activation map for class $k$ at the $i$th timestep as

$$A_i^k = \sum_{c=1}^{C} w_c^k x_i^c \quad (7)$$

so that we have $q_k = \frac{1}{T}\sum_{i=1}^{T} A_i^k$ (an illustration of this calculation can be found at the bottom of Fig. 4). Thus, it is clear that $A_i^k$ can be used to reflect the importance of the $i$th timestep for determining that the input belongs to class $k$. Note that the timestep at the feature map before the GAP layer is reduced by a factor of $2^{N_{\text{sec}}}$ owing to the $N_{\text{sec}}$ max-pooling layers in the model (the pooling size of the max-pooling layer is 2), therefore, the values in the maps at the output are repeated $2^{N_{\text{sec}}}$ times in order to upsample the activation map to the size of the input and obtain the CAM for the input sample.

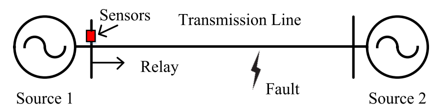



Fig. 5. Studied three-phase power transmission system with sources at both ends. Voltage and current signals on one side of the line are monitored.

# III. Results and Discussion

In this section, the performance of the proposed model is discussed with numerical experiments. Two tasks, namely, fault classification and location, are simultaneously considered as the model target. Section III-A introduces the two datasets utilized in this article. Section III-B provides implementation details of the proposed model. Focusing on dataset I, Section III-C demonstrates the model accuracy and discusses the effect of each modified block on robustness, and Section III-D visualizes the model robustness through explainable components. In Section III-E, the proposed model is adapted to dataset II, and its performance is compared with existing approaches. Finally, in Section III-F, field data are used for verification.

## A. Data Generation

The transmission line system studied in this work is shown in Fig. 5. When a short-circuit fault occurs in the system, the post-fault voltage and current signals are measured by the contactless sensors based on photoelectric [33] and tunnel magnetoresistance [34] at one end of the system. To provide data for the fault classification task that determines the type of fault, the nonfaulty scenario and 10 fault types are considered. Specifically, fault types a–g, b–g, c–g are single-phase-to-ground (L–G) faults, ab, ac, and bc are phase-to-phase (L–L) faults, ab–g, ac–g, and bc–g are two-phase-to-ground (L–L–G) faults, and abc–g is the three-phase-to-ground (L–L–L) fault.

Two datasets from [15] are used to train and test models. Dataset I is generated from a 200 km, 220 kV, 50 Hz transmission system (simulated in MATLAB/Simulink). The transmission line has a positive sequence impedance $Z_1 = 4.76 + j59.75\ \Omega$ and a zero sequence impedance $Z_0 = 77.70 + j204.26\ \Omega$. The impedance at the sources is $Z_s = 3.0135 + j43.35\ \Omega$. The system (simulated in PSCAD/EMTDC) for dataset II is similar to that used for dataset I but has untransposed transmission lines. The impedance of the sources is $Z_s = 9.19 + j74.76\ \Omega$, and the positive and zero sequence impedances of the transmission line are $Z_1 = 5.38 + j84.55\ \Omega$ and $Z_0 = 64.82 + j209.53\ \Omega$, respectively. For both datasets, three-phase voltage and current signals are measured at a sampling frequency of 20 kHz, thus, each cycle has 400 timesteps. The sample length of the datasets is 200 timesteps, leading to an input feature dimension of $6 \times 200$,

TABLE I
SIMULATION PARAMETERS

| System parameter | Dataset | |
|---|---|---|
| | I (MATLAB) | II (PSCAD) |
| Fault distance (km) | randomly generated | randomly generated |
| Fault resistance (Ω) | 0.01, 5, 15, 20 30, 40, 50 | 0.01, 5, 15, 30, 50 |
| Fault inception angle (°) | 0 to 330 (30) | 0 to 300 (60) |
| Prefault power angle (°) | 10, 20, 30 | 10, 20, 30 |

TABLE II
DETAILS OF THE PROPOSED MODEL

| Layer | Description | Output size ($T \times C$) |
|---|---|---|
| $\text{Conv}_1$ | [32 filters, size=3, ReLU] $\times$ 2 | 200 $\times$ 32 |
| $\text{Pool}_1$ | max-pooling, size=2 | 100 $\times$ 32 |
| $\text{Nonlocal}_1$ | $C'$=16 | 100 $\times$ 32 |
| $\text{ResBlock}_2$ | [64 filters, size=3, ReLU] $\times$ 2 | 100 $\times$ 64 |
| $\text{Pool}_2$ | max-pooling, size=2 | 50 $\times$ 64 |
| $\text{Nonlocal}_2$ | $C'$=32 | 50 $\times$ 64 |
| $\text{ResBlock}_3$ | [128 filters, size=3, ReLU] $\times$ 2 | 50 $\times$ 128 |
| $\text{Pool}_3$ | max-pooling, size=2 | 25 $\times$ 128 |
| $\text{Non-local}_3$ | $C'$=64 | 25 $\times$ 128 |

indicating the three-phase current and voltage data over 0.5 cycles.

The parameters used to generate samples in the two datasets are listed in Table I, where the value in the brackets indicates the interval between parameters. For dataset I, a total of 10 different time windows relative to the fault inception time are used. Specifically, the timestep number differences, obtained by subtracting the pertinent number of the fault inception time from that of the starting time, include −80, −60, −40, −20, 0, 20, 40, 60, 100, and 200. Besides, dataset II has the same time window that starts at the fault inception time. Overall, a total of 249 480 samples are simulated for dataset I, while dataset II has 5940 samples. For both datasets, the proportions of the training, validation, and test set are 60%, 20%, and 20%, respectively.

## B. Implementation Details

Details of the shared layers in the proposed multitask model are shown in Table II. The three sections mainly include convolutional layers (or ResBlocks), max-pooling layers, and nonlocal blocks. Moreover, DropBlock layers within $\text{Conv}_1$, $\text{ResBlock}_2$, and $\text{ResBlock}_3$ have $S_b = 11$ and $p_{\text{keep}} = 0.8$, whereas the remaining DropBlock layers have $S_b = 3$ and $p_{\text{keep}} = 0.9$. The hyperparameters of models utilized in this paper are tuned using the grid-search process [35] in order to achieve the best performance.

## C. Accuracy and Robustness Against Data Pollution

The error in the fault location task is given by

$$\text{Error} = \frac{|\text{Output distance} - \text{True distance}|}{\text{Line length}} \times 100\% \tag{8}$$

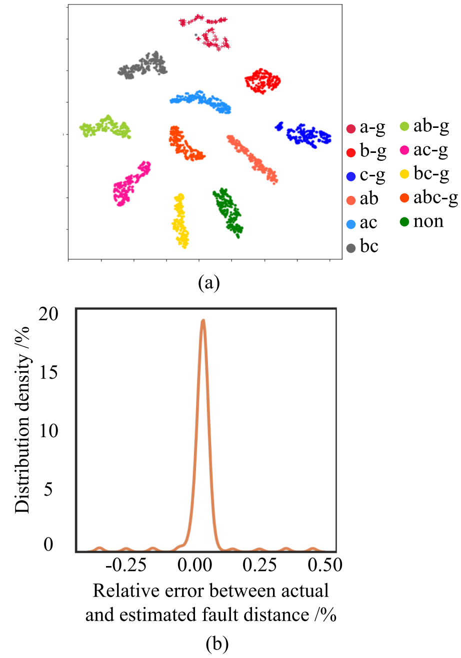



Fig. 6. Performance of two tasks achieved by the proposed integrated structure (in absence of data pollution). (a) t-SNE visualization for fault classification task. (b) relative error distribution in fault location task.

The proposed model is built based on dataset I. In order to prove the effectiveness and high accuracy of the proposed integrated model, the model performance is first obtained in the absence of data pollution. In this case, the accuracy of fault classification is 99.83%, and the error of fault location is 1.98%. For validation, we visualize the feature representations at the GAP layer by t-SNE [36] for fault classification [see Fig. 6(a)], and display the relative error distribution for the fault location task [see Fig. 6(b)]. It is seen that the proposed model effectively distinguishes different fault types and exhibits low errors in fault location estimation.

Robustness of a model refers to its ability to maintain correct results when perturbations including noise and errors appear in the data. In this work, several components are added to the benchmark model to improve the model robustness against data pollution. The model, which was trained by the original simulation data, is validated by the test set with three types of data pollution that are commonly seen in field measurements [15], including the following.

1) *Gaussian noise* that is added to the signals to reach certain signal-to-noise ratios (SNR), with the SNR values being randomly selected from 10, 12, 14, and 16 dB.
2) *Single-timestep errors*, i.e., errors that last for only one timestep. If a signal is affected by this type of error, all its values have a probability $p_{\text{SE}}$ to be replaced by an error value (in this work, 50% of the time it is set to zero, and the other 50% of the time it is set to positive or negative 2, 4, 6, 8, or 10 p.u.) The value of $p_{\text{SE}}$ is randomly chosen from 0.01, 0.02, and 0.03.
3) *Consecutive errors* representing a signal section that takes the same value for more than one timestep. Possible error

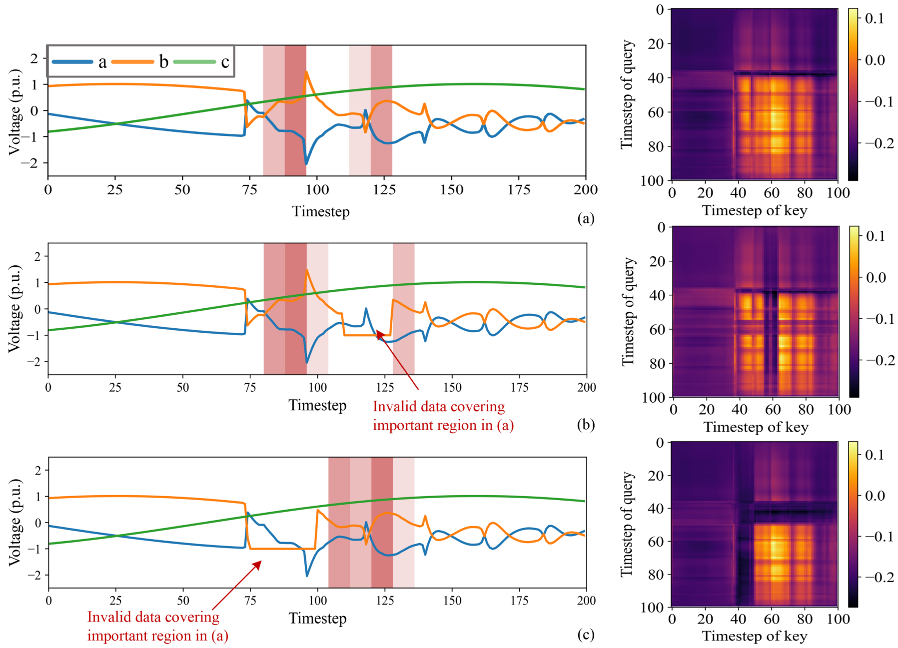



Fig. 7. CAMs of the proposed model (a) without and (b), (c) with data pollution. Left: voltage signals. Right: attention maps of the final nonlocal layer.

TABLE III
PERFORMANCE OF PROPOSED MODEL WITH DIFFERENT COMPONENTS (IN PRESENCE OF DATA POLLUTION)

| Conv | ResBlock | Non-local | DropBlock | Accuracy of classification(%) | Error of location(%) |
|---|---|---|---|---|---|
| ✓ | | | | 92.00 | 4.57 |
| ✓ | ✓ | | | 93.01 | 3.58 |
| ✓ | ✓ | ✓ | | 94.11 | 2.90 |
| ✓ | ✓ | | ✓ | 93.33 | 3.29 |
| ✓ | ✓ | ✓ | ✓ | **98.27** | **2.20** |

lengths include 5, 10, 15, 20, 25, 30, 35, and 40. Candidate choices of amplitude for this error are the same as single-timestep errors. While single-timestep errors may occur more than once, we assume that consecutive errors occur exactly once for the affected signal.

To investigate the contribution of model components on the overall robustness, performances of different model structures are obtained via the removal of specific components, as shown in Table III (where the indexes with the best performance are highlighted in bold hereinafter). It is seen that both the nonlocal block and the DropBlock layer enhance performance of the model, and combining both components further lifts robustness, as the nonlocal block and DropBlock encourage the model to capture global features and information from temporally distributed areas in the input.

Furthermore, the unified structure is characterized by generally higher effectiveness compared to the separate models for fault classification and location tasks. To prove this, individual models for two tasks, which share the same structure as the proposed model [see Fig. 1(b)] except for the output layer, are generated. In the absence of data pollution, the accuracy of fault classification is 99.81%, and the error of fault location is 1.93%, which are close to those using the unified structure. However, in the presence of data pollution, the fault classification accuracy decreases to 97.82% and fault location error increases to 3.73%, implying obvious performance degradation in practical scenarios if separate models are used. Conversely, the unified structure model maintains high performance in both tasks (see Table III). Indeed, since the model input is jointly decided by the fault type and the fault distance in transmission lines, the multitask model can introduce the high relevance between the fault classification and location as an inductive bias [25], and the common knowledge relevant to the two tasks can be shared during the model's optimization process [25], leading to higher robustness against data pollution.

## D. Explainability Analysis

The CAM generated from the GAP layer introduces explainability to the method in the classification task. According to (6) and (7), the CAM highlights the data that contribute most to the model output. When different signals are adopted as inputs to the model, as visualized by CAM and shown in Fig. 7 (the darker region indicates the higher importance hereinafter), the important regions are adaptively determined by model optimization. Specifically, in the first case [see Fig. 7(a)], the input signals are valid throughout the interval, and two high-importance regions are identified for the fault classification task. The pertinent information is focused on by the model, and makes higher contributions to the output.

To reveal the reason behind the high robustness of the proposed model, consecutive invalid data are added to the high-importance regions of the voltage signal in phase B, as shown in Fig. 7(b) and (c), then the most important regions are relocated

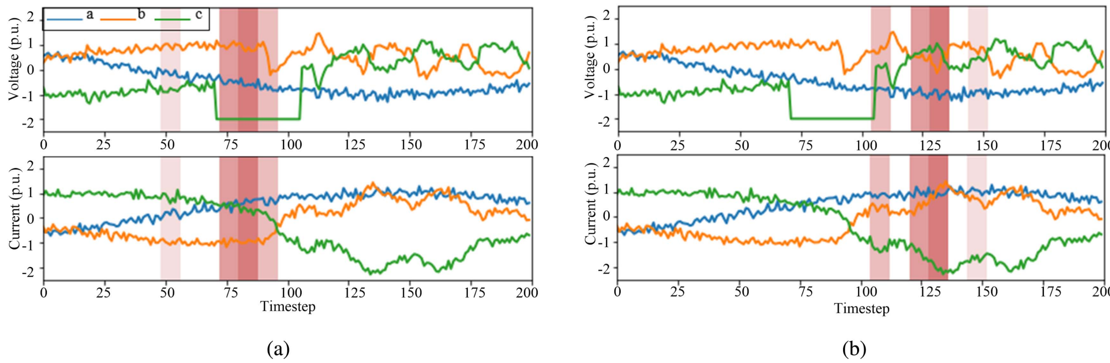



Fig. 8. Example of bc-fault comparing robustness of the benchmark model versus proposed model. (a) Benchmark model. (b) Proposed model.

to the remaining unaffected sections. Thus, the model robustness can be maintained at a high level as the proposed structure adaptively captures valid data. A deeper understanding of this result can be gained by visualizing the attention maps of the final nonlocal layer. The attention scores of each row belong to the same query and are assigned to each timestep according to (2). The attention maps in Fig. 7(b) and (c) neglect the error-affected parts of the signal, yet maintaining attention to the rest of the signal, demonstrating that the model can adaptively focus on valid data.

Then, when a bc-type fault is considered for fault classification task, Fig. 8 visualizes and compares robustness against data pollution between the benchmark and the proposed model. With the benchmark model, three important regions have been placed in the successive error range [see Fig. 8(a)], leading to an incorrect output of fault classification bc–g. Conversely, important regions of the proposed model adaptively avoid the consecutive error data [see Fig. 8(b)], leading to the correct fault classification. Thus, the proposed model structure results in higher robustness.

Despite the efficient use of model explainability in the benchmark and the proposed models, it is noted that explainability is not generally available. Indeed, the benchmark and proposed algorithms contain components that enable the model to exhibit explainability based on CAMs, including the CNN backbone, and the GAP layer. Conversely, the previous models that will be discussed in the following section lack the utilization of these components and, hence, exhibit no explainability.

### E. Comparison Versus Existing Approaches

In this section, the proposed model is built based on dataset II. The model generated based on dataset I can be used as the pretraining model to improve training efficiency. While the proposed model has an integrated framework for fault classification and location tasks, different comparison models, including the deep structure models (LSTM, DSAE, and GSRC) and shallow models (RF and MLP), are trained separately for two tasks. Specifically, for the MLP model proposed in [24], four MLP locators are built for different fault categories (i.e., L–G, L–L, L–L–G, and L–L–L), hence, fault location needs to be estimated after determining the fault type.

TABLE IV
COMPUTATIONAL EFFICIENCY OF DIFFERENT METHODS

| Model | offline processing time/s | | online test time per sample/s | |
|---|---|---|---|---|
| | classification | location | classification | location |
| RF [14] | 61.328 | 65.109 | 0.010 | 0.010 |
| MLP [24] | 59.325 | 52.019 | 0.009 | 0.010 |
| LSTM [21] | 612.817 | 638.948 | 0.192 | 0.228 |
| DSAE [15] | 512.980 | 582.210 | 0.198 | 0.211 |
| GSRC [17] | 479.318 | 504.901 | 0.188 | 0.179 |
| Proposed model | 490.127 | | 0.194 | |

TABLE V
MODEL PERFORMANCE COMPARISON WITH UNPOLLUTED DATA

| Model | Accuracy of classification (%) | Error of location(%) |
|---|---|---|
| RF [14] | 97.46 | 2.09 |
| MLP [24] | 97.10 | 2.15 |
| LSTM [21] | **99.95** | 1.92 |
| DSAE [15] | 98.20 | 1.92 |
| GSRC [17] | 99.94 | 1.91 |
| Proposed model | **99.95** | **1.79** |

The offline training and online testing times of different methods are summarized in Table IV. All computations were executed on a server with a CPU 6240R 2.4 G, 24 C/48 T. As four MLP model is trained in [24] for the fault location task, the average time of the four separate model is taken as the processing time. In general, the RF and MLP models consume less time with their shallow structures, at the expense of their achieved accuracy (as will be discussed below). Among the deep structure models, the proposed model, characterized by a combined structure for accomplishing the two tasks, has training and test times close to those of other models for a single task (see[15], [17], [21]). Hence, the multitask architecture of the proposed model reduces the overall computational burden.

Table V demonstrates the classification accuracies and location errors under unpolluted data. Among the different methods, the shallow models present worse performance than the deep-layer models. Specifically, the MLP model has the worst performance in both tasks, as the simple MLP layers have limited ability to extract high-dimensional electrical data; conversely, the proposed model reaches the highest accuracy in both tasks.

TABLE VI
ACCURACY OF DIFFERENT MODELS UNDER VARIOUS SNR LEVELS FOR FAULT CLASSIFICATION

| Model | SNR (dB) | | | | | |
|---|---|---|---|---|---|---|
| | 20 | 18 | 16 | 14 | 12 | 10 |
| RF [14] | 97.40 | 97.31 | 97.22 | 97.15 | 97.07 | 96.99 |
| MLP [24] | 97.02 | 96.94 | 96.79 | 96.71 | 96.65 | 96.59 |
| LSTM [21] | 99.70 | 99.62 | 99.43 | 99.05 | 98.59 | 98.01 |
| DSAE [15] | 99.72 | 99.61 | 99.47 | 99.21 | 98.64 | 98.01 |
| GSRC [17] | 99.82 | 99.75 | 99.67 | 99.60 | 99.25 | 99.03 |
| Proposed model | **99.83** | **99.82** | **99.78** | **99.78** | **99.70** | **99.67** |

TABLE VII
ERROR OF DIFFERENT MODELS UNDER VARIOUS SNR LEVELS FOR FAULT LOCATION

| Model | SNR (dB) | | | | | |
|---|---|---|---|---|---|---|
| | 20 | 18 | 16 | 14 | 12 | 10 |
| RF[14] | 2.18 | 2.22 | 2.35 | 2.67 | 3.00 | 3.48 |
| MLP[24] | 2.23 | 2.29 | 2.32 | 2.77 | 3.14 | 3.55 |
| LSTM [21] | 1.98 | 2.14 | 2.26 | 2.41 | 2. 88 | 3.31 |
| DSAE [15] | 2.01 | 2.19 | 2.25 | 2.35 | 2.92 | 3.89 |
| GSRC [17] | 1.99 | 2.15 | 2.24 | 2.29 | 2.66 | 2.99 |
| Proposed model | **1.89** | **2.01** | **2.09** | **2.22** | **2.29** | **2.32** |

TABLE VIII
ACCURACY OF DIFFERENT MODELS UNDER VARIOUS ERROR PROPORTIONS FOR FAULT CLASSIFICATION

| Model | Proportion of error (%) | | | | | |
|---|---|---|---|---|---|---|
| | 0.5 | 1 | 1.5 | 2 | 2.5 | 3 |
| RF [14] | 96.00 | 94.08 | 90.94 | 89.20 | 87.17 | 85.38 |
| MLP [24] | 95.97 | 93.02 | 90.89 | 90.09 | 87.25 | 85.61 |
| LSTM [21] | 99.62 | 99.01 | 98.59 | 97.98 | 97.21 | 96.65 |
| DSAE [15] | 97.71 | 95.39 | 93.25 | 92.02 | 90.98 | 90.00 |
| GSRC [17] | **99.88** | 99.50 | 99.02 | 98.56 | 97.69 | 97.03 |
| Proposed model | 99.85 | **99.79** | **99.38** | **99.18** | **99.11** | **99.06** |

TABLE IX
ERROR OF DIFFERENT MODELS UNDER VARIOUS ERROR PROPORTIONS FOR FAULT LOCATION

| Model | Proportion of error (%) | | | | | |
|---|---|---|---|---|---|---|
| | 0.5 | 1 | 1.5 | 2 | 2.5 | 3 |
| RF [14] | 2.39 | 3.19 | 3.99 | 4.69 | 5.98 | 6.87 |
| MLP [24] | 2.39 | 3.06 | 3.88 | 4.59 | 6.00 | 6.93 |
| LSTM [21] | 2.37 | 3.00 | 3.28 | 3.55 | 3.99 | 4.37 |
| DSAE [15] | 2.49 | 3.55 | 4.27 | 4.89 | 5.88 | 6.89 |
| GSRC [17] | 2.36 | 2.96 | 3.18 | 3.50 | 3.82 | 4.29 |
| Proposed model | **2.19** | **2.69** | **3.08** | **3.39** | **3.63** | **3.98** |

Considering the possible presence of various types of data pollution in practice, robustness of the proposed model is further compared with existing approaches, as shown in Tables VI–IX. For better comparison, the data pollution is adjusted based on existing works [15], [17]. Specifically, in the first case, Gaussian noise is added to the test samples, reaching SNRs of 10, 12, 14, 16, 18, and 20. In the second case, consecutive errors are added to the test samples (in one of the input channels), achieving an error proportion between 0.5% and 3%. For 50% of the time, the error amplitude is set to zero, whereas for the rest of the time it is set to a high value (20 p.u.), either positive or negative.

In Tables VI and VII, performances of different models under various SNR levels are listed. While all models are quite robust against noise, the proposed model has the best performance, which can be ascribed to the introduction of nonlocal blocks into the CNN-based backbone. As a matter of fact, this combination of attention computation and Convs has proved robust against Gaussian noise, where the attention operations in nonlocal blocks can be addressed as trainable spatial smoothing of feature maps [37], making the proposed model more robust toward Gaussian noise.

In Tables VIII and IX, the model performances under different proportions of consecutive errors are shown. Among all the methods, LSTM, GSRC, and the proposed model present better performance in both classification and location tasks. Conversely, the performance of RF, MLP, and DSAE degrades rapidly with the increase in consecutive error duration, suggesting that these models are distracted by the continuous invalid information. Meanwhile, in the case of errors with a large amplitude (20 p.u.), the proposed model has the slightest performance degradation compared to LSTM and GSRC, as its nonlocal blocks capture long-range dependencies directly by computing attention interactions between any two positions [27]; the direct dot-product similarity operation (2) in the proposed network adjusts its receptive field to find useful clues and filter out invalid data for fault diagnosis [38]. The same conclusion has been visually proved by the CAMs in Figs. 7 and 8(b), where the model adaptively neglects the consecutive errors and finds related clues to support its output, maintaining correct results under data pollution.

Besides, as previously discussed, the proposed algorithm, owing to its specific components, is characterized by unique explainability compared to other approaches. This helps us to reveal the physical mechanism and prove the model feasibility.

## *F. Verification on Field Data*

In this work, field data measured from the 220 kV Zhanggong-Fuxing transmission line, China [17], are used to validate the proposed model. Three-phase voltage signals were recorded at the Fuxing side of the line with a sampling rate of 4 kHz. The half-cycle signals containing 40 timesteps, with the starting timestep matching the fault inception time, are used as the test sample. As the fault location information is unknown, only the fault classification task is investigated.

Since samples in dataset II were generated using a similar time window that starts at fault inception time, they are downsampled and adjusted (to 40 timesteps that match the field data) to train the proposed model. Afterward, the recorded field data demonstrating an a–g fault are fed into the proposed model, which correctly identifies the fault type [see Fig. 9(a) for three-phase signals and the CAM]. To further validate the proposed model, consecutive errors are added to the single-phase signal [see Fig. 9(b)] and three-phase signals [see Fig. 9(c)] in regions of high importance. In both cases, the model can adaptively relocate the regions of high interest and maintain the correct classification result. This

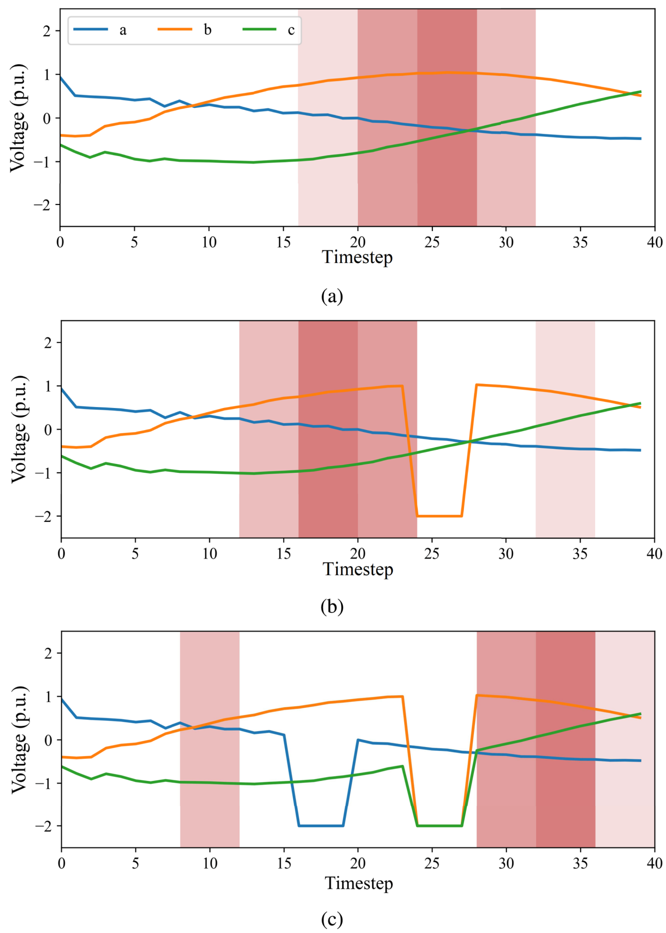



Fig. 9. CAMs of the proposed method with (a) original, and (b), (c) polluted field data.

proves that the proposed model trained by simulation data can be effectively applied to practical cases.

## IV. Conclusion

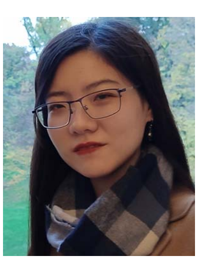
In this article, we proposed a robust and explainable model with a unified structure for fault classification and location on transmission lines. The model robustness against the noise and data errors that are commonly seen in field measurements is verified with numerical experiments. It can be summarized as follows.

1) The proposed method presents a better performance than the existing approaches. In the absence of data pollution, the classification accuracy and location error are 99.83% and 1.98%, respectively. In the presence of data pollution, the model maintains high robustness owing to its unified structure and adopted components.
2) The visualization using CAMs and attention maps reveals that the proposed model can adaptively neglect consecutive errors in signals and focus on the valid information, intuitively explaining the reason behind model robustness, and proving the feasibility of the data-driven method in practical applications.
3) Since the model explainability and robustness are gained by processing data in specific examples, the proposed model has the potential to be applied to other fault diagnosis tasks with similar characteristics, such as fault detection and location of the electric power cables that are widely installed in cities.

## References

[1] D. Chen, Z. Xia, L. Wei, X. Feng, and Z. D. Wang, “A survey of methods for improving the UHV power system transfer capacity,” in *Proc. IEEE Int. Conf. Sustain. Power Gener. Supply*, 2012, pp. 1–4.
[2] M. Singh, B. K. Panigrahi, and R. P. Maheshwari, “Transmission line fault detection and classification,” in *Proc. IEEE Int. Conf. Emerg. Trends Elect. Comput. Technol.*, 2011, pp. 15–22.
[3] K. Chen, Z. He, S. X. Wang, J. Hu, L. Li, and J. He, “Learning-based data analytics: Moving towards transparent power grids,” *CSEE J. Power Energy Syst.*, vol. 4, no. 1, pp. 67–82, 2018.
[4] D. Novosel, G. Bartok, G. Henneberg, P. Mysore, and S. Ward, “IEEE PSRC report on performance of relaying during wide-area stressed conditions,” *IEEE Trans. Power Del.*, vol. 25, no. 1, pp. 3–16, Jan. 2010.
[5] H. R. Salim, K. C. O. Salim, and S. A. Bretas, “Further improvements on impedance-based fault location for power distribution systems,” *IET Generation, Transmiss. Distrib.*, vol. 5, no. 4, pp. 467–478, Apr. 2011.
[6] Y. Xi, Z. Li, X. Zeng, X. Tang, X. Zhang, and H. Xiao, “Fault location based on travelling wave identification using an adaptive extended Kalman filter,” *IET Gener., Transmiss. Distrib.*, vol. 12, no. 6, pp. 1314–1322, 2018.
[7] V. C. Gungor, B. Lu, and G. P. Hancke, “Opportunities and challenges of wireless sensor networks in smart grid,” *IEEE Trans. Ind. Electron.*, vol. 57, no. 10, pp. 3557–3564, Oct. 2010.
[8] A. Jamehbozorg and S. Shahrtash, “A decision tree-based method for fault classification in double-circuit transmission lines,” *IEEE Trans. Power Del.*, vol. 25, no. 4, pp. 2184–2189, Oct. 2010.
[9] H. Livani and C. Y. Evrenosoglu, “A fault classification and localization method for three-terminal circuits using machine learning,” *IEEE Trans. Power Del.*, vol. 28, no. 4, pp. 2282–2290, Oct. 2013.
[10] K. Silva, B. Souza, and N. Brito, “Fault detection and classification in transmission lines based on wavelet transform and ANN,” *IEEE Trans. Power Del.*, vol. 21, no. 4, pp. 2058–2063, Oct. 2006.
[11] Y. Q. Chen, O. Fink, and G. Sansavini, “Combined fault location and classification for power transmission lines fault diagnosis with integrated feature extraction,” *IEEE Trans. Ind. Electron.*, vol. 65, no. 1, pp. 561–569, Jan. 2017.
[12] A. Pradhan, A. Routray, S. Pati, and D. Pradhan, “Wavelet fuzzy combined approach for fault classification of a series-compensated transmission line,” *IEEE Trans. Power Del.*, vol. 19, no. 4, pp. 1612–1618, Oct. 2004.
[13] P. Dash, S. Samantaray, and G. Panda, “Fault classification and section identification of an advanced series-compensated transmission line using support vector machine,” *IEEE Trans. Power Del.*, vol. 22, no. 1, pp. 67–73, Jan. 2007.
[14] D. Patil, O. D. Naidu, P. Yalla, and S. Hida, “An ensemble machine learning based fault classification method for faults during power swing,” in *Proc. IEEE Innov. Smart Grid Technol. - Asia*, 2019, pp. 4225–4230.
[15] K. Chen, J. Hu, and J. He, “Detection and classification of transmission line faults based on unsupervised feature learning and convolutional sparse autoencoder,” *IEEE Trans. Smart Grid*, vol. 9, no. 3, pp. 1748–1758, May 2018.
[16] J. Wu, “Introduction to convolutional neural networks,” *Nat. Key Lab Novel Softw. Technol.*, Nanjing Univ., Nanjing, China, vol. 5, 2017, Art. no. 23.
[17] S. Shi, B. Zhu, S. Mirsaeidi, and X. Dong, “Fault classification for transmission lines based on group sparse representation,” *IEEE Trans. Smart Grid*, vol. 10, no. 4, pp. 4673–4682, Jul. 2019.
[18] X. Qin et al., “A cable fault recognition method based on a deep belief network,” *Comput. Elect. Eng.*, vol. 71, pp. 452–464, 2018.
[19] X. Lei and Z. Sui, “Intelligent fault detection of high voltage line based on the faster R-CNN,” *Measurement*, vol. 138, pp. 379–385, 2019.
[20] J. Liang, T. Jing, H. Niu, and J. Wang, “Two—terminal fault location method of distribution network based on adaptive convolution neural network,” *IEEE Access*, vol. 8, pp. 54035–54043, 2020.

[21] S. Belagoune, N. Bali, A. Bakdi, B. Baadji, and K. Atif, “Deep learning through LSTM classification and regression for transmission line fault detection, diagnosis and location in large-scale multi-machine power systems,” *Measurement*, vol. 177, no. 3, 2021, Art. no. 109330.
[22] L. Zhu and D. J. Hill, “Cost-effective bad synchrophasor data detection based on unsupervised time-series data analytic,” *IEEE Internet Things J.*, vol. 8, no. 3, pp. 2027–2039, Feb. 2021.
[23] P. Ray and D. P. Mishra, “Support vector machine based fault classification and location of a long transmission line,” *Eng. Sci. Technol. Int. J.*, vol. 19, no. 3, pp. 1368–1380, 2016.
[24] A. K. Onaolapo, R. P. Carpanen, D. G. Dorrell, and E. E. Ojo, “Transmission line fault classification and location using multi-layer perceptron artificial neural network,” in *Proc. IEEE 46th Annu. Conf. Ind. Electron. Soc.*, 2020, pp. 5182–5187.
[25] A. Kumar and H. D. Iii, “Learning task grouping and overlap in multi-task learning,” *Comput. Sci.*, vol. 2, pp. 1–8, 2012.
[26] Y. Zhou, Q. Guo, H. Sun, Z. Yu, J. Wu, and L. Hao, “A novel data-driven approach for transient stability prediction of power systems considering the operational variability,” *Int. J. Elect. Power Energy Syst.*, vol. 107, pp. 379–394, 2019.
[27] X. Wang, R. Girshick, A. Gupta, and K. He, “Non-local neural networks,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 7794–7803.
[28] G. Ghiasi, T.-Y. Lin, and Q. V. Le, “DropBlock: A regularization method for convolutional networks,” in *Proc. Adv. Neural Inf. Process. Syst.*, 2018, pp. 10727–10737.
[29] A. Jafar and L. Myungho, “Hyperparameter optimization for deep residual learning in image classification,” in *Proc. IEEE Int. Conf. Autonomic Comput. Self-Organizing Syst. Companion*, 2020, pp. 24–29.
[30] S. Lan, M.-J. Chen, and D.-Y. Chen, “A novel double-terminal non-synchronous fault location method based on convolutional neural network,” *IEEE Trans. Power Del.*, vol. 34, no. 3, pp. 848–857, Jun. 2019.
[31] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” in *Proc. 3rd Int. Conf. Learn. Representations*, 2015, pp. 1–14.
[32] B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, and A. Torralba, “Learning deep features for discriminative localization,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 2921–2929.
[33] Y. X. Nie, X. G. Yin, L. I. Wei, and D. X. Huang, “Design of DC power supply for active photoelectric transducer,” *Elect. Power Autom. Equip.*, vol. 23, no. 12, pp. 49–51, 2003.
[34] L. Xue, H. Wei, D. Song, and W. Jian, “Design and verification of the current transducer based on the tunneling magneto-resistance effect,” in *Proc. 2nd Int. Conf. Electric Inf. Control Eng.*, 2012, vol. 3, pp. 278–281.
[35] A. C. Florea and R. Andonie, “Weighted random search for hyperparameter optimization,” *Int. J. Comput., Commun. Control*, vol. 14, no. 2, pp. 154–169, 2019.
[36] L. van der Maaten and G. Hinton, “Visualizing data using t-SNE,” *J. Mach. Learn. Res.*, vol. 9, pp. 2579–2605, 2008.
[37] N. Park and S. Kim, “How do vision transformers work?,” 2022. [Online]. Available: https://arxiv.org/abs/2202.06709
[38] M. M. Naseer, K. Ranasinghe, S. H. Khan, M. Hayat, F. Shahbaz Khan, and M.-H. Yang, “Intriguing properties of vision transformers,” in *Proc. Adv. Neural Inf. Process. Syst.*, 2021, vol. 34, pp. 23296–23308.



**Jiashu Fang** received the B.S. degree in electrical engineering from North China Electric Power University, Beijing, China, in 2015 and the M.S. degree in electrical power system engineering from the University of Manchester, Manchester, U.K., in 2016. She is currently working toward the Ph.D. degree in electrical engineering with North China Electric Power University.

Her research interests include applications of machine learning and data science in power systems.

**Kunjin Chen** received the B.Sc. and Ph.D. degrees in electrical engineering from Tsinghua University, Beijing, China, in 2015 and 2020, respectively.

He currently a Machine Learning Expert with Alibaba Group, Hangzhou, China. His research interests include applications of machine learning in smart grid and supply chains, especially in the field of time series forecasting.

**Chongru Liu** (Senior Member, IEEE) received the B.S. and Ph.D. degrees in electrical engineering from Tsinghua University, Beijing, China, in 2000 and 2006, respectively.

She is currently a Professor, Ph.D. Supervisor, and the Dean of Undergraduate Academic Affairs Office, North China Electric Power University, Beijing, China. Beijing Science and Technology Nova Talent Support Plan was selected in 2009, and the Ministry of Education Excellent Talents Support Plan for the New Century in 2012.

Dr. Liu is or was the Standing Director of the Tenth Council of Beijing Electrical Engineering Society and the Committee on Science Popularization and Education, Member of the Fifth and Sixth National Technical Committee on Power System Management and Information Exchange Standardization, Power System Automation Committee of China Electrical Engineering Society, Foreign Affairs Committee of China Electrical Engineering Society, and Beijing Science Representatives of the Ninth Congress of the Technical Association.

**Jinliang He** (Fellow, IEEE) received the B.Sc. degree from the Wuhan University of Hydraulic and Electrical Engineering, Wuhan, China, in 1988, the M.Sc. degree from Chongqing University, Chongqing, China, in 1991, and the Ph.D. degree from Tsinghua University, Beijing, China, in 1994, all in electrical engineering.

He became a Lecturer in 1994, and an Associate Professor in 1996, with the Department of Electrical Engineering, Tsinghua University. From 1997 to 1998, he was a Visiting Scientist with Korea Electrotechnology Research Institute, Changwon, South Korea. From 2014 to 2015, he was a Visiting Professor with the Department of Electrical Engineering, Stanford University, Palo Alto, CA, USA. In 2001, he was promoted to a Professor with Tsinghua University, where he is currently the Chair with High Voltage Research Institute. He has authored seven books and 500 technical papers. His research interests include advanced power transmission technology, sensing technology and big data mining, and smart nanodielectric materials.