Renewable Energy and Power Quality Journal
https://repqj.com/index.php/repqj/index
RE&PQJ, Vol. 23, No. 1, 2025

Application of Fusion Algorithm of Insulator Partial Discharge Acoustic Characteristics and Image Identification in Electricity Transmission Line Detection

Guojian Shan*, Han Wang, Yanpeng Wang

Transmission Work Area, State Geid Xingan Electric Power Supply Company, Xinganmeng 137400, Neimenggu, China
*Corresponding author's email: 15149032040@163.com

Abstract. Traditional electricity transmission line detection methods usually rely on a single technical means and lack multi-dimensional comprehensive diagnosis, resulting in insufficient accuracy and comprehensiveness in fault identification. This paper constructs an electricity transmission line detection model based on the insulator partial discharge acoustic characteristics and image identification technology to improve the accuracy and intelligence level of fault detection. During model building, the acoustic characteristics and image data of insulator partial discharge are first obtained through the equipment. After data preprocessing, it is input into the convolutional neural networks (CNN) and recurrent neural networks (RNN) for training. Through parameter optimization, the model shows high efficiency and accuracy in electricity transmission line fault detection. Experimental results show that the model exhibits good efficiency and accuracy in electricity transmission line fault detection through parameter adjustment. According to the experimental results, the detection accuracy of the proposed model on various data sets is 93.40%, 91.50%, and 89.20%, respectively, better than that of other control models, and the processing time is 118 seconds, 182 seconds, and 238 seconds, respectively, also better than that of control groups. In conclusion, the model constructed in this paper provides a reliable and effective method for electricity transmission line fault detection and lays a foundation for the advancement of related technologies.

Key words. Insulator Partial Discharge, Acoustic Characteristics, Image Identification, Fusion Algorithm, Electricity Transmission Line Detection

1. Introduction

Transmission line faults seriously affect the power supply stability and safety of contemporary power systems [1,2]. In the case of continuously increasing power demand, the operating environment of transmission lines becomes increasingly complex, with increased types and frequency of faults. Although infrared imaging, temperature monitoring, and vibration analysis have been traditional methods of fault detection widely used in early fault diagnosis, they have an obvious limitation in the handling of diversified faults in complicated environments. It is only hot spots that the infrared imaging methods can identify via the temperature distribution on the surface of the transmission lines. However, its shortcomings lie in it being able to monitor only the change in surface temperature and cannot easily provide effective diagnosis for deep-seated or potential faults. All the more, under high-temperature or high-humidity environments, the accuracy of temperature distribution will be disturbed and lead to a misjudgment [3]. Temperature monitoring is mainly dependent on the layout of sensors. Although it can realize the real-time monitoring of equipment temperature changes, the density of the temperature sensor is normally not enough for the full coverage of all transmission lines' potential failure points, while its sensitivity against slight failures or localized problems is so low, failing to afford thorough fault identification [4]. The vibration analysis method collects the mechanical vibration signals and then deduces whether any mechanical faults occurred. However, this method is susceptible to interference from external vibration sources, especially in windy conditions. It is difficult to effectively distinguish between external vibration and equipment fault signals, reducing the accuracy of fault diagnosis [5]. Besides, the traditional methods have poor ability in identifying fault types, especially under complicated or variable operating conditions. It is difficult to deal with hidden problems such as micro discharges or potential faults that are difficult to detect. Therefore, although these traditional methods have been used in early diagnosis, they are still insufficient in terms of accuracy, timeliness and adaptability of fault detection, and cannot meet the needs of modern power systems for efficient and intelligent fault detection [6,7]. As a typical

fault precursor, the acoustic characteristics of partial discharge have great potential in early fault diagnosis [8,9]. The high-frequency acoustic wave signals generated by partial discharge can provide effective early warning through monitoring. However, fault detection based on acoustic signals alone often reduces the accuracy. First, partial discharge signals are of high frequency and small amplitude, which is prone to the effect of electromagnetic interference and mechanical vibration noise. In particular, complex environments can cause these noises to be mistaken for partial discharge signals. Second, under different fault types, the performance of partial discharge signals varies greatly; hence, one acoustic feature can hardly distinguish various faults with precision. In the process of propagation, it will also be transformed into atmospheric changes of temperature, humidity, and wind speed, which will further reduce the accuracy of diagnosis. Accordingly, in such complex environments, the limitation of a single method of acoustic detection exists. With a view to improve the accuracy of fault detection, researchers have started to explore the method of multimodal data fusion combined with other sensor data, including image, temperature, and vibration signals, which can improve the detection performance to a certain extent in terms of accuracy and robustness [10,11].

This paper integrates the detection scheme of partial discharge acoustic characteristics of insulators with image identification technology, overcomes the limitations of traditional single technology, and improves the accuracy and intelligence level of fault detection through multi-modal data fusion. This paper will discuss how partial discharge acoustic signals can be combined with image identification technology for efficient and comprehensive fault detection in transmission lines. Besides promoting the development of smart power monitoring technology and providing technical support for the reliable operation of future power systems, this research also provides a new method for the detection of faults in electricity transmission lines.

This paper is innovative, especially in its unique multimodal data fusion strategy. First, in terms of data fusion, a weighted fusion method is designed to effectively complement the acoustic signal and image features, maximizing their respective advantages. It superimposes the time-domain characteristics of the acoustic signals with the spatial information of the image data based on this, and comprehensively expresses the fault characteristic more informatively to further enhance the distinguishability of the model concerning the type of fault. In particular, a deep learning algorithm has been applied for joint learning on the multiple sensors in this work to optimize the level of intelligence on feature extraction and fault identification. By this fusion model, the detected system can make automatic identifications and classify the fault type in a far more effective way, greatly improving the intellectualization level in fault detection and meeting the requirements of intelligent power monitoring technology developments.

Main contributions: This study proposes a multimodal deep neural network method that combines the acoustic characteristics of partial discharge of insulators with image recognition, which significantly improves the accuracy and robustness of transmission line fault detection. The specific innovations are as follows: a weighted fusion method is used to dynamically adjust the weight coefficient, optimize feature contribution, reduce modal interference, and enhance the recognition ability of key features. Combining the multimodal architecture of CNN and RNN, image spatial features and acoustic temporal features are extracted respectively to achieve comprehensive fault detection.

Comparative experiments verify the advantages of the fusion model in different scale data sets, noise conditions and detection speed. Under 2000, 5000 and 8000 data sets, the detection accuracy of the experimental group is 93.40%, 91.50% and 89.20% respectively, the recall rate is 91.10%, 89.20% and 86.50% respectively, and the processing time is 118 seconds, 182 seconds and 238 seconds respectively, which are better than the control group. The actual application evaluation showed that the experimental group was ahead of the control group in terms of user satisfaction, detection accuracy and ease of use. The specific evaluation results were: ease of use 4.6 points, detection speed 4.8 points, detection accuracy 4.9 points, usage satisfaction 4.7 points, and ease of operation 4.6 points.

2. Related Work

Recent studies have shown that the application of traditional detection methods in modern complex power systems is significantly limited, especially in changing environments and high load conditions, where the accuracy and stability of traditional methods are greatly reduced [12,13]. For example, Jalil B pointed out that infrared imaging technology has poor stability in extreme climates and is easily disturbed by environmental factors [14]. Zainuddin N M believes that temperature monitoring methods cannot cope with temperature changes caused by load fluctuations or equipment aging, resulting in potential faults being difficult to accurately identify [15]. Vibration analysis requires a larger fault amplitude to be effectively detected, and early or minor faults are difficult to capture in time [16]. These limitations make it difficult for traditional methods to meet the needs of real-time and comprehensive fault detection in complex power systems. In contrast, detection methods based on machine learning can improve the accuracy and timeliness of fault detection by integrating multiple data sources. Machine learning algorithms can process high-dimensional and noisy data and automatically identify potential fault modes, thereby overcoming the shortcomings of traditional methods. This makes the application of machine learning in complex power systems, especially in early fault detection and prediction, show great advantages.

With the rapid development of intelligent technology,

machine learning, image recognition and sound signal analysis have been widely used in transmission line defect detection [17-19]. Among them, CNN is used for automatic detection of insulator cracks and corrosion due to its advantages in feature extraction and image processing, and can accurately capture key fault characteristics [20]. RNN and its variants (such as LSTM) perform well in analyzing the time series data of partial discharge acoustic signals. By learning time series patterns, early diagnosis of electrical faults can be achieved [21]. In addition, the combination of CNN and RNN further enhances the multimodal data processing capability. By jointly analyzing images and acoustic signals, it not only reduces the impact of noise interference, but also improves the recognition accuracy and intelligence level of complex faults.

Research on the fusion of images and acoustic signals has attracted attention in recent years [22,23]. Multimodal fusion uses the physical fault information provided by images and the electrical fault signs revealed by acoustic signals to achieve more comprehensive fault detection. However, existing fusion methods face challenges in information extraction and integration [24,25]. First, the modal differences between images and acoustic signals make feature extraction and alignment difficult. Second, traditional weighted fusion methods cannot dynamically adjust feature weights, which easily leads to the loss of important information. Finally, the robustness and real-time performance of existing fusion algorithms in complex environments are insufficient. To solve these problems, research can introduce a dynamic feature fusion framework based on deep learning, such as using attention mechanisms and multimodal contrastive learning to improve feature alignment and weight allocation accuracy. In addition, optimizing the network structure and adopting lightweight models can improve real-time performance.

3. Construction of Electricity Transmission Line Detection Model

A. Overview of Research Framework

1) Model Design Concept

This paper's electricity transmission line detection model overcomes the drawbacks of a single technology, combines image identification technology with the acoustic characteristics of partial discharge, and achieves more precise fault diagnosis. The acoustic signals of partial discharge can detect electrical faults at an early stage, especially when there is no apparent external damage. Image identification processes images of insulator surfaces to identify physical faults such as cracks and corrosion. The fusion model improves the precision and intelligence level of fault detection through the complementary effect of the two data sources. Figure 1 shows the specific design process:

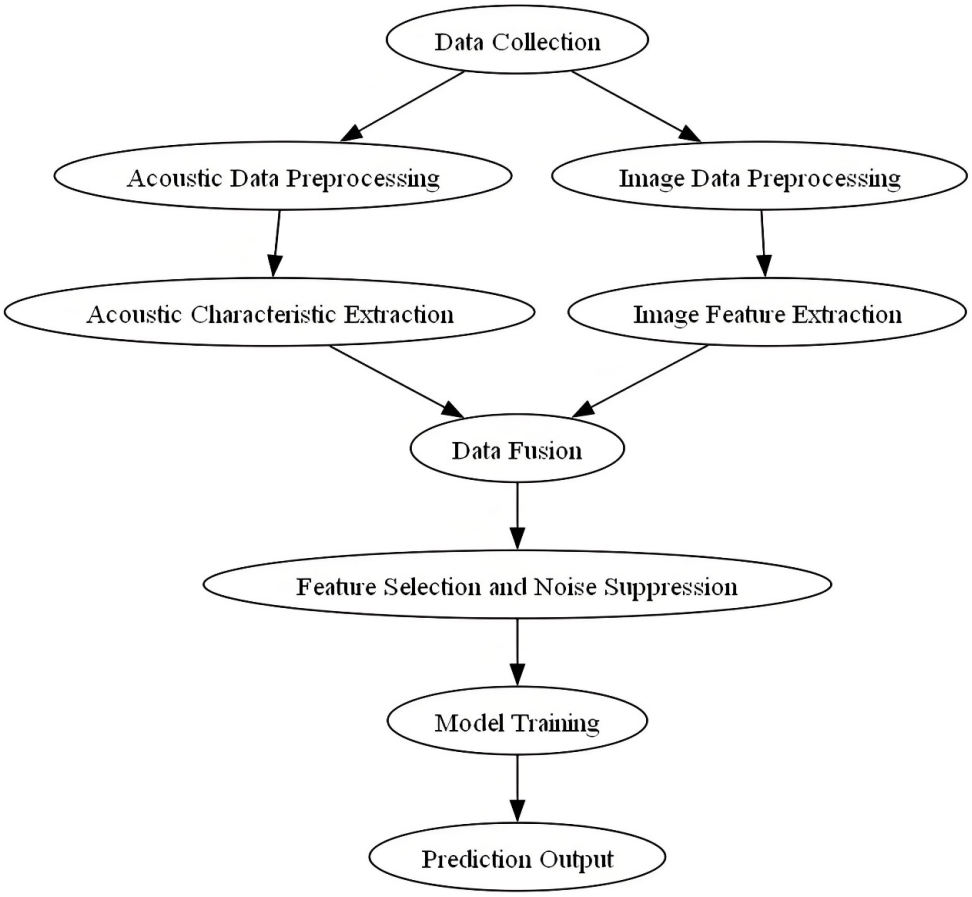

Figure 1. Design process of the fusion model

Specifically, the acoustic and image data are first preprocessed to extract their key features respectively. Next, the acoustic characteristics and image features are integrated into the model for training using a multi-modal neural network architecture. In model design, the key is to choose a suitable fusion strategy to give full play to the advantages of the two types of data. In addition, the model needs to consider the possible heterogeneity and noise interference of different data sources. Therefore, in the data fusion stage, it is necessary to apply effective noise suppression and feature selection methods to improve the model's stability and robustness.

2) Advantages of Multi-Modal Data

The core advantage of the multi-modal data fusion model is that it integrates complementary information from different data sources and overcomes the limitations of a single data source. In electricity transmission line fault detection, acoustic characteristics can sensitively reflect electrical faults, especially in detecting early discharges and minor electrical faults. Image data can effectively identify physical faults such as cracks and corrosion on the surface of insulators. Combining acoustic signals and image data enables the system to provide a more comprehensive and precise diagnosis under various fault types.

During the fusion process, image data helps identify potential problems other than electrical faults, such as physical faults such as mechanical damage, corrosion, and cracks. These problems may appear before electrical faults occur, and image data can provide early visual information to help the detection system identify these hidden dangers. Acoustic signals can warn of potential electrical faults through high-frequency sound waves when electrical faults do not appear as physical damage. Through multimodal data fusion, the system can extract information from a wider feature space and improve the accuracy of fault diagnosis, especially in complex power environments, enhancing the robustness and reliability of detection.

B. Data Collection and Preprocessing

1) Data Collection Process

In this paper, two main devices are used for data collection. The partial discharge acoustic signals are collected by high-sensitivity acoustic sensors and industrial acoustic imagers to capture the discharge phenomena on the insulator's surface precisely. Image data is acquired through high-resolution digital cameras and infrared thermal imagers. The former is used to collect physical damage, such as cracks and corrosion, and the latter is used to detect abnormal overheating.

During the data collection process, 3,000 sets of samples are collected. Each data set includes a complete monitoring cycle, covering acoustic signals and image data in different operating environments. Each monitoring cycle lasts 5 minutes, and acoustic and image data are collected synchronously to ensure the fault features are fully reflected. The collected data is stored for subsequent processing and model training, covering changing conditions such as temperature and humidity to enhance data diversity and the model's generalization ability.

2) Acoustic Characteristic Extraction

This study combines frequency domain and time domain analysis methods to extract the acoustic characteristics of partial discharge signals to fully utilize the advantages of both. In time domain analysis, the waveform and amplitude changes of the signal can reveal the temporal characteristics of partial discharge events, which is especially important for the identification of early faults and signal mutations. Frequency domain analysis can provide spectral information of the signal to help capture high-frequency components, which are usually closely related to the occurrence of electrical faults. By combining time domain and frequency domain analysis, the temporal and frequency characteristics of partial discharge signals can be fully reflected, thereby improving the accuracy of fault detection, especially for accurate diagnosis of early faults in complex environments.

This paper adopts the statistical feature extraction method in the time domain analysis. The basic characteristics of partial discharge signals are described by calculating the signal's peak value, mean value, standard deviation, variance, and other statistical quantities. Specifically, the peak value reflects the maximum amplitude of the signal:

$$\text{Peak} = \max\left(x(t)\right) \quad (1)$$

In Formula (1), $x(t)$ is a time domain signal. The mean and variance help describe the stability and volatility of the signal. Their calculation formulas are as follows:

$$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i \quad (2)$$

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}\left(x_i - \mu\right)^2 \quad (3)$$

The Fourier transform (FT) is adopted in this paper to perform spectral analysis on the data, converting the time domain signal into frequency domain information, thereby helping to identify the main frequency components and frequency changes in the signal [26]. The formula for the Fourier transform is:

$$X(f)=\int_{-\infty}^{\infty}x(t)\mathrm{e}^{-j2\pi ft}\mathrm{d}t \quad (4)$$

The frequency features of partial discharge signals have significant regularity, and the frequency distribution and energy distribution of different discharge signals vary greatly. Therefore, different fault types can be distinguished by extracting spectrum features such as main frequency, bandwidth, and frequency peak. The extraction formula for main frequency (f_{main}) is as follows:

$$f_{\text{main}}=\text{argmax}\left|X(f)\right| \quad (5)$$

The bandwidth is calculated as follows:

$$BW=f_{\max}-f_{\min} \quad (6)$$

In Formula (6), $f_{\max}$ and $f_{\min}$ are the maximum and minimum frequencies in the frequency distribution, respectively. Additionally, this paper applies the time-frequency analysis method and uses wavelet transform to process non-stationary signals. The formula of continuous wavelet transform is:

$$W(a,b)=\int_{-\infty}^{\infty}x(t)\psi^{*}\left(\frac{t-b}{a}\right)\mathrm{d}t \quad (7)$$

Finally, after the above analysis, this paper extracts the following key acoustic characteristics: discharge frequency, amplitude, discharge duration, transient change spectrum, signal energy distribution, etc. These characteristics provide an essential basis for subsequent fault diagnosis, especially in identifying early electrical faults, with high accuracy and sensitivity.

3) Image Preprocessing

Image preprocessing is the first step in image data analysis, which aims to improve image quality, highlight fault features, and remove noise interference. First, the collected color image is grayscaled to convert into a grayscale image to simplify subsequent analysis [27]. Grayscale is achieved through the following formula:

$$I_{\text{gray}}=0.2989\cdot I_{\text{R}}+0.5870\cdot I_{\text{G}}+0.1140\cdot I_{\text{B}} \quad (8)$$

In Formula (8), I_{R}, I_{G}, and I_{B} are the pixel values of the image's red, green, and blue channels, respectively. After grayscale conversion, image processing is more efficient and convenient for subsequent feature extraction.

This paper chooses the Sobel operator for image edge detection, mainly based on its good balance between computational efficiency and noise suppression. The Sobel operator calculates the gradient of the image in the horizontal and vertical directions, which can not only highlight the detailed features of the edge, but also has a certain smoothing effect, which can effectively reduce the impact of noise on the edge detection results. In addition, compared with more complex edge detection algorithms, the Sobel operator is simple to calculate and has high efficiency, which is very suitable for processing large-scale image data [28]. The Sobel operator retrieves an image's edge features by calculating each pixel's gradient. It uses two 3×3 convolution kernels to calculate the gradient of the image in the horizontal and vertical dimensions. The following are formulas for calculating gradients:

$$G_x=\begin{bmatrix}-1 & 0 & 1\\ -2 & 0 & 2\\ -1 & 0 & 1\end{bmatrix} \quad (9)$$

$$G_y=\begin{bmatrix}-1 & -2 & -1\\ 0 & 0 & 0\\ 1 & 2 & 1\end{bmatrix} \quad (10)$$

Through the convolution operation, the Sobel operator calculates the horizontal and vertical gradients of each pixel to obtain a gradient image, thereby extracting the edge information of the image, especially structures such as cracks and damage. This paper uses the histogram equalization method to more clearly present the tiny cracks or damage on the insulator's surface to enhance the image's contrast and details. Considering the noise interference, this paper uses Gaussian filtering to denoise the image to further improve the image quality.

4) Data Fusion Preparation

In the data fusion preparation stage, this study adopted a precise synchronization method based on timestamps to ensure that acoustic data and image data are collected and effectively matched within the same time window. Specifically, during the data acquisition process, all sensing devices are calibrated by a unified high-precision time synchronization module to ensure that the time reference of the acoustic sensor and the image acquisition device is consistent. In actual operation, a high-precision timestamp is attached to each frame of image data, and the acoustic signal is sampled in real time and segmented. Each acoustic signal is associated with a corresponding timestamp. Subsequently, by matching the timestamps of the image data and the acoustic signal, a one-to-one correspondence between the two is constructed to ensure that each frame of image data accurately corresponds to the complete acoustic signal within the acquisition period. This method effectively avoids synchronization errors caused by equipment time deviation or sampling frequency differences, and provides a reliable basic guarantee for subsequent data fusion.

Then, this paper applies the Z-score normalization method to eliminate the dimensionality differences in the data. This method transforms the data into a regular normal distribution with a mean of zero and a standard deviation of one to compare and fuse various dimensional data on the same scale. Specifically, the mean and standard deviation of the image and acoustic data are determined respectively, and then a normalization procedure is applied:

$$Z = \frac{x - \mu}{\sigma} \quad (11)$$

In Formula (11), the raw data is represented by x. The data mean is represented by μ. The data standard deviation is represented by σ. The normalized data is represented by Z. The fused data is divided into a training set and a validation set in a ratio of 8:2. When dividing the data, the number of each type of fault should be balanced to avoid adverse effects of data imbalance on model training.

C. Electricity Transmission Line Detection Model Architecture and Training

1) Model Architecture Design

The electricity transmission line fault detection model designed in this paper adopts a multi-modal deep neural network (DNN) architecture that combines CNN and recurrent neural networks (RNN) [29,30]. Figure 2 presents its specific architecture:

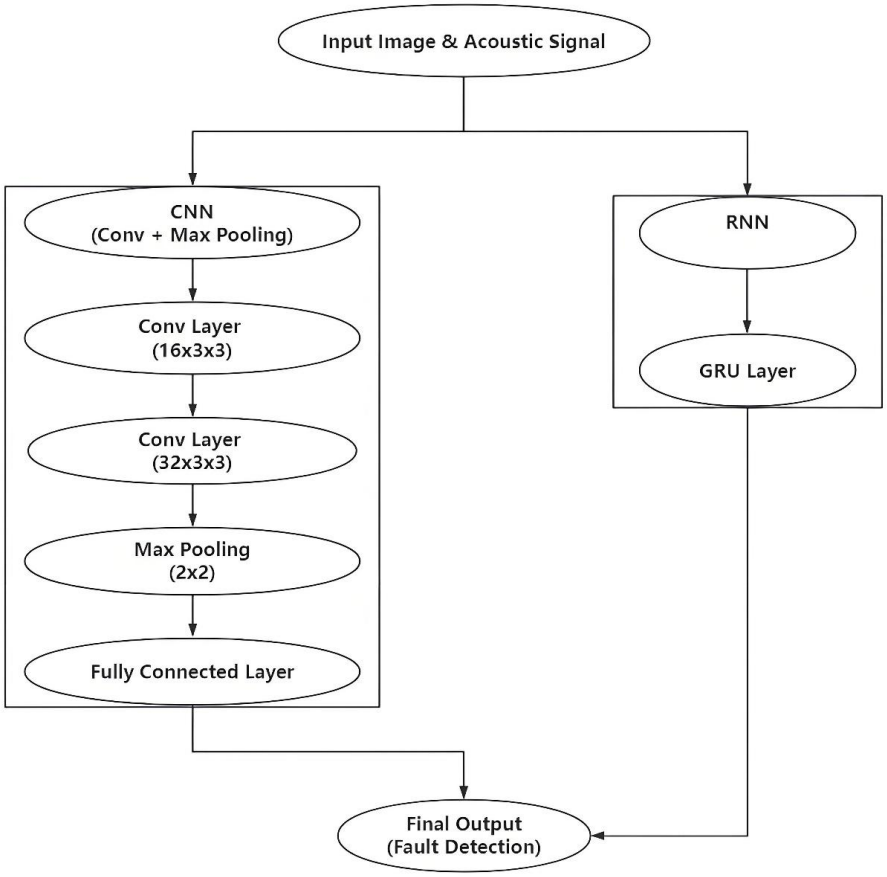

Figure 2. Structure of the electricity transmission line detection model

The CNN module is mainly used to extract local details and structural features from images and is the preferred architecture for this task. This is because CNN can efficiently capture local patterns such as edges and textures of images through convolution operations, and has parameter sharing and local connection characteristics, which greatly reduces computational complexity. Compared with traditional machine learning algorithms, CNN does not need to rely on manually designed features, but automatically learns low-level to high-level feature representations through hierarchical feature extraction, thereby capturing image information more comprehensively and significantly improving the accuracy and robustness of fault detection. In this study, the CNN module consists of multiple convolutional layers, pooling layers, and fully connected layers. The first convolutional layer contains 16 convolution kernels with a convolution kernel size of 3×3, a stride of 1, and 1-pixel zero padding (padding = 1), and uses the ReLU activation function to capture basic edge and texture features. Subsequently, the second and third convolutional layers contain 32 and 64 convolution kernels, respectively, and also use 3×3 convolution kernels and ReLU activation functions to extract more complex structural features. The formula for the convolution operation is:

$$y(i,j) = (f * x)(i,j) = \sum_m \sum_n f(m,n)\, x(i+m, j+n) \quad (12)$$

In Formula (12), $f(m,n)$ is the convolution kernel, and $y(i,j)$ is the output feature map after the convolution operation.

The output of each convolutional layer is then passed through a maximum pooling layer for feature dimensionality reduction, where the pooling window is 2×2 and the stride is 2. The maximum pooling operation can be expressed by the following formula:

$$y(i,j) = \max_{m,n} x(i+m, j+n) \quad (13)$$

After convolution and pooling, the feature map is flattened and integrated into high-dimensional image features through a fully connected layer, helping the model learn the complex structure of the image at a higher level.

The RNN module processes acoustic signal data and is particularly good at capturing dynamic changes in time series data. This study uses a gated recurrent unit (GRU) structure to overcome the gradient vanishing problem of traditional RNN in long sequence processing. GRU dynamically controls the flow of information by updating the gate and resetting the gate, selectively retaining or forgetting information, thereby effectively capturing the time dependency and long-term characteristics of the local discharge signal. In the specific implementation, each acoustic signal is represented as a time series and input into the GRU module. The GRU module consists of two hidden layers, each containing 128 hidden units, and uses the Tanh activation function for learning nonlinear features.

2) Feature Fusion Method

Fusing acoustic characteristics and image features is a key step to improve the fault detection model's performance. This paper adopts a weighted fusion method to give full play to the complementary advantages of acoustic signals and image data. In this method, different weights are assigned to different features to ensure that each feature can play the greatest role in subsequent learning according to its importance. Specifically, the acoustic characteristic F_{ac} and the image feature F_{img} are first processed and extracted, respectively, and then weighted summed according to predetermined weight coefficients α and β $(\alpha + \beta = 1)$ to form a fused feature vector F_{fused}. The specific calculation formula is as follows:

$$F_{fused} = \alpha \cdot F_{ac} + \beta \cdot F_{img} \quad (14)$$

In the weighted fusion process, the dynamic adjustment of weight coefficients α and β is based on the change of loss value in the model training feedback. Specifically, after each round of training, the fusion model will calculate the contribution of each modal feature to the final output based on the feedback of the loss function. When the feature of a certain modality contributes more in the current training round, its corresponding weight coefficient will be slightly increased, while the weight of the modality with less contribution will be appropriately reduced. The adjustment process can be achieved through the normalized gradient weighting method, the formula is as follows:

$$\alpha_{t+1} = \frac{\partial L}{\partial \alpha_t} \cdot \eta,\ \beta_{t+1} = \frac{\partial L}{\partial \beta_t} \cdot \eta \quad (15)$$

Through dynamic adjustment, the weight coefficients α and β can effectively adapt to the impact of acoustic signals and image features on model performance at different stages, ensuring the balance of the fusion process in data scale and information expression, thereby improving the accuracy and robustness of fault detection.

3) Model Training and Optimization

In model training, this paper adopts a supervised learning method and trains based on the data set processed in the previous paper. The core goal of training is to optimize the model's parameters by minimizing the loss function so that it can accurately predict electricity transmission line faults under different input conditions. Figure 3 shows the training process:

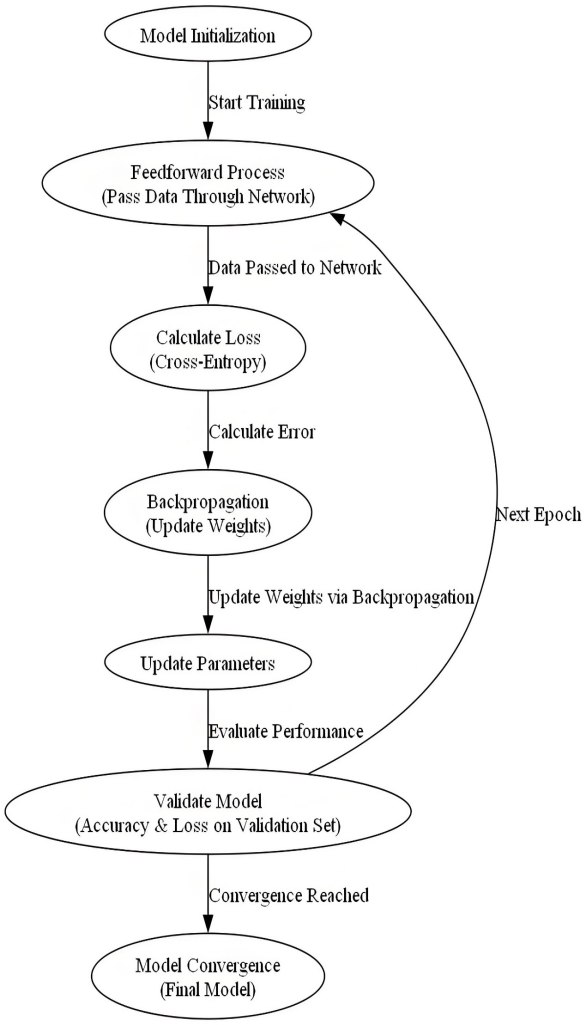

Figure 3. Training process of the electricity transmission line detection model

The training process starts with the initialization of the model, and the number of training epochs is 50. The data is passed to each network layer through the feedforward process to generate preliminary prediction results. Subsequently, the error between the predicted value and the actual value is calculated using the cross-entropy loss function, and the error is passed back to each network layer through the backpropagation algorithm to update the model weights. After each training epoch, the model updates the parameters according to the training data, and the model performance under the current hyperparameter configuration is evaluated through the validation set. Through multiple epochs of iterations, the model gradually reduces the prediction error and eventually converges.

Since the task of this paper is a binary classification problem, the cross-entropy loss function is selected as the loss function. A popular indicator for classification problems is the cross-entropy loss function. It determines the overall loss by summing the log difference between the true label and the expected probability value for each sample. During training, the model gradually improves parameters by minimizing the cross-entropy loss to ensure continuous improvement in performance on the training set.

This paper uses L2 regularization and dropout techniques to improve the model's generalization ability and prevent overfitting. L2 regularization reduces model complexity and the risk of overfitting by adding a penalty term of squared weights to the loss function. The dropout technique enhances model robustness and avoids over-reliance on specific nodes by randomly discarding some neurons. At the same time, this paper adopts the Adam optimization algorithm to accelerate training and avoid gradient disappearance or explosion. The algorithm combines the momentum method and adaptive learning rate to dynamically adjust the learning rate of each parameter, thereby accelerating the convergence process. After each training epoch, the model is evaluated on the validation set, using accuracy and training loss as indicators. The learning rate and other hyperparameters are adjusted based on the evaluation results to ensure the stability of the model on different data sets.

In order to improve the efficiency and accuracy of the model, this study used the grid search method to optimize the hyperparameters. The specific steps include:

first, the adjustment range of the two key hyperparameters, learning rate and batch size, is defined. The adjustment range of the learning rate is from 10^{-5} to 10^{-1}, and the selection range of the batch size is 32, 64, and 128. Then, the grid search method is used to perform an exhaustive search in these preset hyperparameter spaces to evaluate the impact of each set of hyperparameter combinations on model performance.

The grid search systematically traverses all possible hyperparameter combinations, combines cross-validation to evaluate the performance of the model, and finally selects the best hyperparameters. After many experiments, the learning rate 10^{-4} and batch size 64 that are most suitable for this study were determined, thus ensuring efficient convergence of the model and the best training effect.Figure 4 presents the training loss and accuracy changes of the model with this configuration:

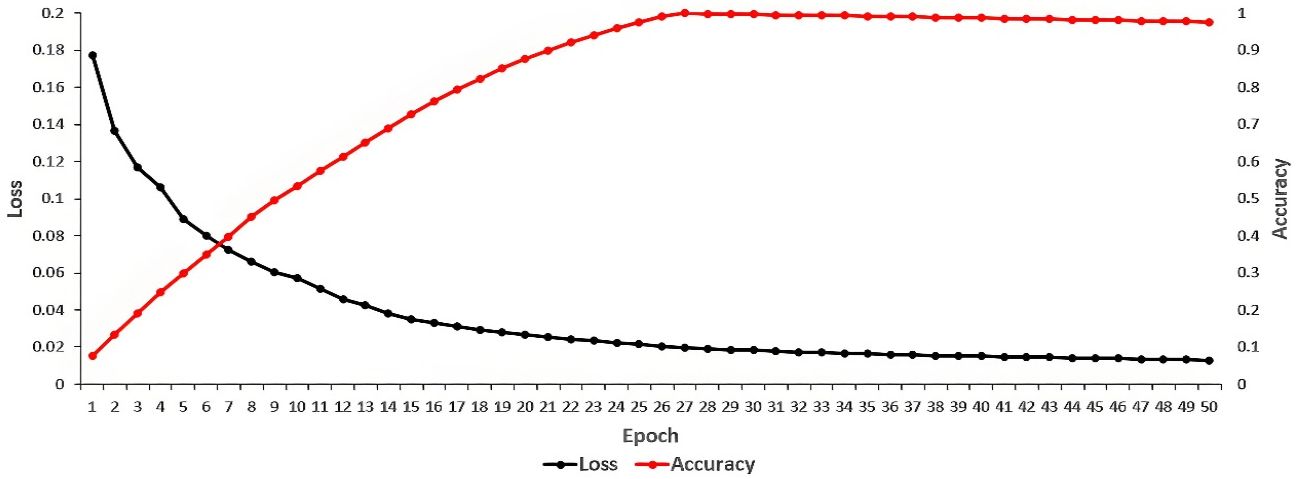

Figure 4. Model training loss and accuracy changes

In Figure 4, the model's accuracy is close to 1 after 50 training epochs, indicating that the model achieves the best performance with this configuration and the accuracy stabilizes. As the training progresses, the loss value gradually decreases, indicating that the model is continuously optimizing and progressively converging. The training process is in line with the expected convergence trend. The changes further verify the effective training of the model with the optimal hyperparameter configuration, and the model's fit continues to improve with the increase in iterations.

4. Experimental Design and Results

A. Experimental Environment and Design

The experimental environment of this paper mainly relies on a high-performance computing platform. The operating system is Ubuntu 18.04. The deep learning framework is TensorFlow 2.0. The Python version is 3.7. All experimental data and models are stored on SSD hard drives to reduce delays in data reading and model training and ensure efficient operation of the experiments. In the experimental environment, the external device interface uses standard communication protocols to achieve real-time connection with sensors, cameras and data acquisition terminals. The interface design includes data reception, protocol analysis and data transmission modules, and the timestamp synchronization mechanism is used to accurately align multi-source data to ensure data integrity and timeliness.In addition, the experiment uses CUDA version 10.2 to support graphics processing unit (GPU) acceleration, which significantly improves the efficiency of the training and inference processes.

In terms of experimental data, 15,000 sets of data are re-collected and preprocessed to construct the experimental data set of this paper to avoid overfitting. In addition, this paper conducts a control experiment and sets up four control groups, namely electricity transmission line detection models based on support vector machine (SVM) (control group 1), random forest (RF) (control group 2), single CNN (control group 3), and single RNN (control group 4), to thoroughly verify the performance of the constructed detection model. After training each control model, they are deployed using the same configuration as the model in this paper to avoid other factors affecting the experimental results.

B. Model Detection Effect and Efficiency Evaluation

2,000, 5,000, and 8,000 sets of data are extracted from the experimental data set, respectively, to form small, medium, and large-scale data sets to evaluate the actual detection effect of the model. The experimental group and control group models are used for testing. During the experiment, the detection accuracy, recall, and processing time of each group of models are recorded. Table 1 lists the detection accuracy and recall of each group:

Table 1. Model detection effect experiment

Data Set Size	Group	Accuracy	Recall
2,000	Experimental Group	93.40%	91.10%
	Control Group 1 (SVM)	86.70%	83.60%
	Control Group 2 (RF)	85.50%	83.00%
	Control Group 3 (CNN)	90.20%	88.30%
	Control Group 4 (RNN)	89.70%	87.80%
5,000	Experimental Group	91.50%	89.20%
	Control Group 1 (SVM)	84.30%	81.80%
	Control Group 2 (RF)	83.60%	80.90%
	Control Group 3 (CNN)	88.00%	86.10%
	Control Group 4 (RNN)	87.50%	85.40%
8,000	Experimental Group	89.20%	86.50%
	Control Group 1 (SVM)	82.10%	79.40%
	Control Group 2 (RF)	80.30%	78.00%
	Control Group 3 (CNN)	85.60%	83.20%
	Control Group 4 (RNN)	84.70%	82.50%

According to the data in Table 1, under data sets of different sizes, there are differences in the performance of the experimental group and the four control groups regarding detection accuracy and recall. First, the experimental group performs well in all data set sizes, with its accuracy and recall being higher than those of control groups in each data set size. Specifically, in the 2,000 data sets, the accuracy of the experimental group is 93.40%, and the recall is 91.10%, significantly better than 86.70% and 83.60% of the control group 1 (SVM), 85.50% and 83.00% of the control group 2 (RF), 90.20% and 88.30% of the control group 3 (CNN), and 89.70% and 87.80% of the control group 4 (RNN). Similarly, as the size of the data set increases, the experimental group's performance remains ahead. In the 5,000 and 8,000 data sets, the accuracy and recall of the experimental group are 91.50% and 89.20%, and 89.20% and 86.50%, respectively, both better than other control groups.

Specifically, in the comparison with the control groups, the performance of the SVM and RF models is relatively poor in all data set sizes, especially in large-scale data sets, where both the accuracy and recall do not exceed 85%. Although the CNN and RNN models perform better at certain scales, especially in small-scale data sets, with the CNN model achieving an accuracy of 90.20%, they still fail to surpass the experimental group on large-scale data sets, and their recall rates are also slightly inferior.

Figure 5 shows the detection time of each group under different data sizes:

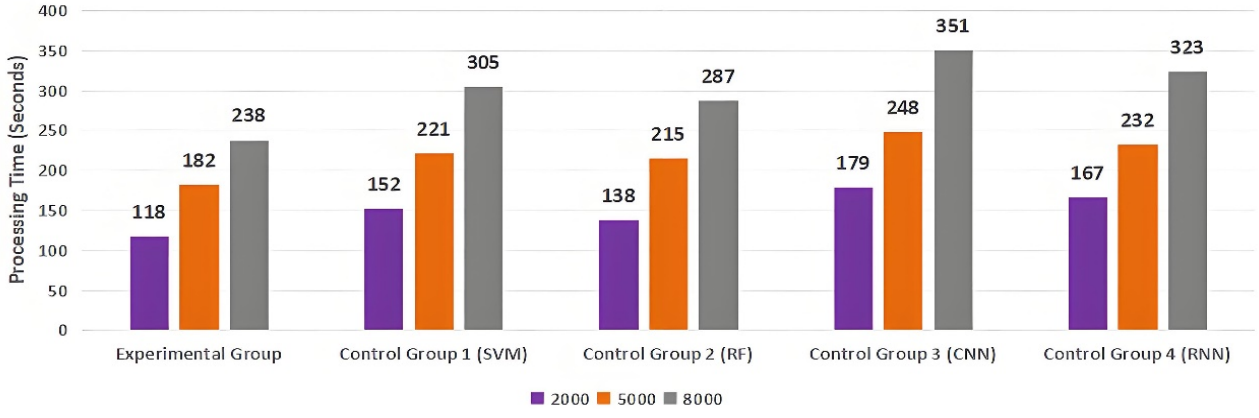

Figure 5. Results of the detection time of each group

Results in Figure 5 show that under data sets of different scales, the processing efficiency of the experimental group is significantly better than that of the control groups. Specifically, when the data scale is 2,000, the detection time of the experimental group is 118 seconds, which is significantly lower than the detection time of each model in the control groups. Especially in control group 3 (CNN) and control group 4 (RNN), the detection time is 179 seconds and 167 seconds, respectively, with significant gaps. Although the detection time of control group 1 (SVM) and control group 2 (RF) is slightly shorter, it is still significantly higher than that of the

experimental group. As the data size increases, the processing efficiency of the experimental group still maintains an advantage. For the data sets of 5,000 and 8,000, the detection time of the experimental group is 182 seconds and 238 seconds, respectively. Although the processing time of each group generally increases with the increase in data volume, the experimental group always maintains the lowest detection time. In particular, compared with control group 3 (CNN) and control group 4 (RNN), the time gaps of the experimental group further widen, indicating that the processing efficiency of the experimental group under large-scale data is still superior.

C. Model Robustness Evaluation

Three subsets of 5,000 data sets are randomly extracted from the experimental data set to evaluate the robustness of the model. In each extraction, different levels of noise (low, medium, high) are added. Then, the three sets of data are input into each model for detection, and each set's detection accuracy and recall are recorded. Table 2 shows the detection results:

Table 2. Experimental results of model robustness

Noise Level	Group	Accuracy (%)	Recall (%)
Low-level Noise	Experimental Group	92.3	90.5
	Control Group 1 (SVM)	84.5	81.7
	Control Group 2 (RF)	83.2	80.4
	Control Group 3 (CNN)	89.1	87.2
	Control Group 4 (RNN)	88.5	86.8
Medium-level Noise	Experimental Group	89.1	86.3
	Control Group 1 (SVM)	81.6	77.8
	Control Group 2 (RF)	80.5	76.6
	Control Group 3 (CNN)	85.5	83.1
	Control Group 4 (RNN)	84.2	82
High-level Noise	Experimental Group	85.8	83
	Control Group 1 (SVM)	75.9	72.3
	Control Group 2 (RF)	74.2	70.5
	Control Group 3 (CNN)	80.1	78.2
	Control Group 4 (RNN)	78.9	76.5

The data in Table 2 shows that under different noise conditions, the robustness of the experimental group is always better than that of other control groups, especially when facing medium-level and high-level noise, its detection effect is still excellent. Under low-level noise conditions, the experimental group achieves an accuracy of 92.3% and a recall of 90.5%, which is the best performance, much higher than all methods in the control groups. Although the performance of each group is not much different, the experimental group still shows stable high performance. In a medium-level noise environment, although the performance of all models declines, the accuracy of the experimental group is 89.1%, and the recall is 86.3%, which remains at a high level. In contrast, the accuracy of the control group 1 (SVM) and the control group 2 (RF) performs poorly, dropping to 81.6% and 80.5%, respectively, and the recall also drops significantly, indicating that these models are more sensitive to noise. Although the control group 3 (CNN) and the control group 4 (RNN) perform better under medium-level noise, they are still inferior to the experimental group. The experimental group generally performs significantly better than other control groups under low-level, medium-level, and high-level noise conditions. Especially when the noise intensity increases, the decrease in its accuracy and recall rate is relatively small, demonstrating its strongest robustness.

D. Users' Actual Application Evaluation

In order to evaluate the effect of the model in practical application, this study invited 50 power maintenance personnel from a power company to participate in the experiment. In the experiment, the participants used the experimental group and control group models to detect faults on the transmission lines. After the detection, questionnaires were distributed to the participants to collect their evaluation of the experience of using each model. The questionnaire design includes multiple evaluation dimensions, such as the ease of operation of the model, fault detection accuracy, response time, model stability, etc., aiming to fully reflect the subjective feelings of maintenance personnel on the performance of the model. Participants scored each indicator using a Likert scale (1-5 points), with 1 indicating very dissatisfied and 5 indicating very satisfied. In addition, suggestions and opinions from participants on model improvements were collected. Finally, the results of the collected questionnaires were statistically analyzed to obtain the evaluation data, as shown in Table 3 below:

Table 3. Results of users' evaluation

Evaluation Dimensions (Out of 5)	Experimental Group	Control Group 1 (SVM)	Control Group 2 (RF)	Control Group 3 (CNN)	Control Group 4 (RNN)
Ease of Use	4.6	4.1	4	4.3	4.2
Detection Speed	4.8	4.3	4.2	4.4	4.5
Detection Accuracy	4.9	4.2	4.1	4.4	4.3
Satisfaction with Use	4.7	4	3.9	4.2	4.1
Ease of Operation	4.6	4.2	4.1	4.3	4.2

According to the data in Table 3, the participants' evaluation of each model shows that the experimental group is better than control groups in all evaluation dimensions. Regarding ease of use, the experimental group scores 4.6, significantly higher than other models, indicating that its interface and interaction design are more in line with user needs and provide a better user experience. Regarding detection speed, the experimental group scores 4.8, better than all control groups, reflecting its efficient real-time fault detection capability. Regarding detection accuracy, the experimental group scores 4.9, significantly better than other models, indicating that it has apparent advantages in identifying electricity transmission line fault features. Regarding satisfaction with use, the experimental group scores 4.7, reflecting the participants' high recognition of its overall performance. Finally, the experimental group scores 4.6 in ease of operation, ahead of the control groups, indicating that its operation is more intuitive and easy to use, reducing the difficulty of user operation.

5. Conclusion

This paper constructs an electricity transmission line detection method based on the fusion of insulator partial discharge acoustic characteristics and image identification technology. A multi-modal deep neural network architecture is designed by combining partial discharge acoustic signals with image data, effectively improving the accuracy and efficiency of electricity transmission line fault detection. This method adopts CNN to extract spatial features in images and uses RNN to analyze the temporal characteristics of acoustic signals, thereby achieving efficient monitoring of electricity transmission lines.

Experimental results show that the fusion algorithm performs well in multiple data sets and different noise environments. Especially in high-level noise conditions, its detection accuracy and robustness are better than traditional single models. In tests on data sets of different sizes, the experimental group also significantly outperforms the control groups in detection speed and accuracy. By comparing the application effects of different models, the experimental group shows higher ease of use, detection speed, and accuracy, and the users are more satisfied with the use, proving the advantages of this method in practical applications.

Although this study has achieved certain results, there are still limitations. First, model training is highly dependent on labeled data, which may increase the cost of data preparation in practical applications. Future research can focus on exploring semi-supervised learning, self-supervised learning, or weakly supervised learning methods to reduce the demand for labeled data by utilizing a large amount of unlabeled data or introducing generative models. Secondly, although this study showed good robustness under a variety of noise conditions, its performance in extremely complex environments still needs further verification and optimization. Future research plans to conduct experiments in more practical application scenarios, and design more targeted optimization strategies for complex environments such as multi-source interference or dynamically changing scenarios, so as to further improve the applicability and stability of the method.

Acknowledgment

None

Data Availability Statement

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

Funding

None

Author Contribution

Guojian Shan: Conceived and designed the research, conducted experiments, and analyzed data. Drafted and revised the manuscript critically for important intellectual content.

Han Wang: Contributed to the acquisition, analysis, and interpretation of data. Provided substantial intellectual input during the drafting and revision of the manuscript.

Yanpeng Wang: Participated in the conception and design of the study.

All authors have read and approved the final version of the manuscript.

Conflicts of Interest

The authors declare that they have no financial conflicts of interest.

References

References

[1] M. Ghiasi, N. Ghadimi, E. Ahmadinia. An analytical methodology for reliability assessment and failure analysis in distributed power system. SN Applied Sciences. 2019, 1(1), 44-53. DOI: 10.1007/s42452-018-0049-0

[2] A.S. Altaie, A.A. Majeed, M. Abderrahim, A. Alkhazraji. Fault detection on power transmission line based on wavelet transform and scalogram image analysis. Energies. 2023, 16(23), 7914-7933. DOI: 10.3390/en16237914

[3] L. Yang, J.F. Fan, Y.H. Liu, E. Li, J.Z. Peng, et al. A review on state-of-the-art power line inspection techniques. IEEE Transactions on Instrumentation and Measurement. 2020, 69(12), 9350-9365. DOI: 10.1109/TIM.2020.3031194

[4] K. Chen, Y. Yue, Y.J. Tang. Research on temperature monitoring method of cable on 10 kV railway power transmission lines based on distributed temperature sensor. Energies. 2021, 14(12), 3705-3720. DOI: 10.3390/en14123705

[5] X.B. Huang, Y.Q. Wu, Y. Zhang, B.T. Li. Structural defect detection technology of transmission line damper based on UAV image. IEEE Transactions on Instrumentation and Measurement. 2022, 72, 1-14. DOI: 10.1109/TIM.2022.3228008

[6] F.M. Shakiba, S.M. Azizi, M.C. Zhou, A. Abusorrah. Application of machine learning methods in fault detection and classification of power transmission lines: a survey. Artificial Intelligence Review. 2023, 56(7), 5799-5836. DOI: 10.1007/s10462-022-10296-0

[7] A. Raza, A. Benrabah, T. Alquthami, M. Akmal. A review of fault diagnosing methods in power transmission systems. Applied Sciences. 2020, 10(4), 1312-1339. DOI: 10.3390/app10041312

[8] H.D. Ilkhechi, M.H. Samimi. Applications of the acoustic method in partial discharge measurement: A review. IEEE Transactions on Dielectrics and Electrical Insulation. 2021, 28(1), 42-51. DOI: 10.1109/TDEI.2020.008985

[9] K. Maresch, L.F. Freitas-Gutierres, A.L. Oliveira, A.S. Borin, G. Cardoso, et al. Advanced Diagnostic Approach for High-Voltage Insulators: Analyzing Partial Discharges through Zero-Crossing Rate and Fundamental Frequency Estimation of Acoustic Raw Data. Energies. 2023, 16(16), 6033-6054. DOI: 10.3390/en16166033

[10] J.M. Liu, Z.Y. Zhao, J. Ji, M.L. Hu. Research and application of wireless sensor network technology in power transmission and distribution system. Intelligent and Converged Networks. 2020, 1(2), 199-220. DOI: 10.23919/ICN.2020.0016

[11] C.M. Furse, M. Kafal, R. Razzaghi, Y.J. Shin. Fault diagnosis for electrical systems and power networks: A review. IEEE Sensors Journal. 2020, 21(2), 888-906. DOI: 10.1109/JSEN.2020.2987321

[12] W.X. Ding, X.Y. Jing, Z. Yan, L.T. Yang. A survey on data fusion in internet of things: Towards secure and privacy-preserving fusion. Information Fusion. 2019, 51, 129-144. DOI: 10.1016/j.inffus.2018.12.001

[13] H. Choi, J.P. Yun, B.J. Kim, H. Jang, S.W. Kim. Attention-based multimodal image feature fusion module for transmission line detection. IEEE Transactions on Industrial Informatics. 2022, 18(11), 7686-7695. DOI: 10.1109/TII.2022.3147833

[14] B. Jalil, G.R. Leone, M. Martinelli, D. Moroni, M.A. Pascali, et al. Fault detection in power equipment via an unmanned aerial system using multi modal data. Sensors. 2019, 19(13), 3014-3029. DOI: 10.3390/s19133014

[15] N.M. Zainuddin, M.S.A. Rahman, M.Z.A.A. Kadir, N.H. Nik Ali, Z. Ali, et al. Review of thermal stress and condition monitoring technologies for overhead transmission lines: Issues and challenges. IEEE Access. 2020, 8, 120053-120081. DOI: 10.1109/ACCESS.2020.3004578

[16] L. Zhao, X.B. Huang, Y. Zhang, Y. Tian, Y. Zhao, et al. A vibration-based structural health monitoring system for transmission line towers. Electronics. 2019, 8(5), 515-526. DOI: 10.3390/electronics8050515

[17] I. Maduako, C.F. Igwe, J.E. Abah, O.E. Onwuasoanya, G.A. Chukwu, et al. Deep learning for component fault detection in electricity transmission lines. Journal of Big Data. 2022, 9(1), 81-115. DOI: 10.21203/rs.3.rs-1028973/v1

[18] W.Q. Zhao, M.F. Xu, X.F. Cheng, Z.B. Zhao. An insulator in transmission lines recognition and fault detection model based on improved faster RCNN. IEEE Transactions on Instrumentation and Measurement. 2021, 70, 1-8. DOI: 10.1109/TIM.2021.3112227

[19] Y. Dash, A. Abraham, N. Kumar, M. Raj. Role of artificial intelligence in transmission line protection: A review of three decades of research. International Journal of Hybrid Intelligent Systems. 2024, 20(3), 185-206. DOI: 10.3233/HIS-240016

[20] H.G. Liang, C. Zuo, W.M. Wei. Detection and evaluation method of transmission line defects based on deep learning. IEEE Access. 2020, 8, 38448-38458. DOI: 10.1109/ACCESS.2020.2974798

[21] X.Y. Zheng, R. Jia, L.L. Gong, G.R. Zhang, J. Dang, et al. Component identification and defect detection in transmission lines based on deep learning. Journal of Intelligent & Fuzzy Systems. 2021, 40(2), 3147-3158. DOI: 10.3233/JIFS-189353

[22] Y.H. Xi, W.J. Zhang, F. Zhou, X. Tang, Z.W. Li, et al. Transmission line fault detection and classification based on SA-MobileNetV3. Energy Reports. 2023, 9, 955-968. DOI: 10.1016/j.egyr.2022.12.043

[23] Z.X. Cai, T.J. Wang, W.Y. Han, A.A. Ding. PGE-YOLO: A Multi-Fault-Detection Method for Transmission Lines Based on Cross-Scale Feature Fusion. Electronics. 2024, 13(14), 2738-2756. DOI: 10.3390/electronics13142738

[24] A. Swain, E. Abdellatif, A. Mousa, P.W.T. Pong. Sensor technologies for transmission and distribution systems: a review of the latest developments. Energies. 2022, 15(19), 7339-7376. DOI: 10.3390/en15197339

[25] A.S. Altaie, M. Abderrahim, A.A. Alkhazraji. Transmission Line Fault Classification Based on the Combination of Scaled Wavelet Scalograms and CNNs Using a One-Side Sensor for Data Collection. Sensors. 2024, 24(7), 2124-2147. DOI: 10.3390/s24072124

[26] X.G. Li, L.X. Yuan, H. Song. Time-Frequency domain conversion of UHVDC transmission line Corona current based on Fourier transform. Electric Power Components and Systems. 2022, 50(13), 696-705. DOI: 10.1080/15325008.2022.2139438

[27] X.B. Huang, T.T. Nie, Y. Zhang, X.L. Zhang. Study on hydrophobicity detection of composite insulators of transmission lines by image analysis. IEEE Access. 2019,

7, 84516-84523. DOI: 10.1109/ACCESS.2019.2922279
[28] Z.B. Qiu, X. Zhu, C.B. Liao, D.Z. Shi, W.Q. Qu. Detection of transmission line insulator defects based on an improved lightweight YOLOv4 model. Applied Sciences. 2022, 12(3), 1207-1224. DOI: 10.3390/app12031207
[29] Y.J. Zhai, X. Yang, Q.M. Wang, Z.B. Zhao, W.Q. Zhao. Hybrid knowledge R-CNN for transmission line multifitting detection. IEEE Transactions on Instrumentation and Measurement. 2021, 70, 1-12. DOI: 10.1109/TIM.2021.3096600
[30] C. Ji, F. Zhang, X.B. Huang, Z.W. Song, W. Hou, et al. STAE-YOLO: Intelligent detection algorithm for risk management of construction machinery intrusion on transmission lines based on visual perception. IET Generation, Transmission & Distribution. 2024, 18(3), 542-567. DOI: 10.1049/gtd2.13093