ingegneria sismica
International Journal of Earthquake Engineering
Anno XLIII - Num. 2 - 2026

Research on image recognition technology of power equipment based on deep residual network

Xiaolong Chen[1], *Chao Sun*[1,*], *Xuming Ni*[1], *Xiaoyuan Jia*[1], *Xing Wang*[1] *and Jian Pang*[1]

[1] Guangdong Power Grid Co., Ltd. Guangzhou Power Supply Bureau, Guangzhou, Guangdong, 510000, China

SUMMARY: *Addressing the issues of weak distinguishability of thermal defect features and a large number of model parameters in infrared image diagnosis of power equipment, this study constructed an infrared image dataset for substation equipment and proposed a thermal defect diagnosis method for power equipment based on deep residual networks. First, convolutional kernel decomposition technology was used to simplify the basic structure of the network, significantly reducing the parameter size of the model. A multi-scale convolutional feature fusion strategy is then employed to integrate semantic features from both shallow and deep layers of the network, thereby enhancing the diagnostic accuracy of thermal defect states. Finally, a Bayesian optimization algorithm based on coupled constraints is designed to adaptively adjust hyperparameters such as the number of convolutional kernels and network depth, enabling lightweight identification. Experiments show that the thermal defect recognition accuracy of this model can reach 93.12% in a simple background, and the optimized thermal defect diagnosis model for power transformation equipment can effectively classify seven different types of thermal defects. This method provides a reliable technical route for intelligent diagnosis of power equipment.*

KEYWORDS: *power transformation equipment; infrared images; thermal defect diagnosis; Bayesian optimization*

1 Introduction

Currently, power equipment inspection is primarily carried out through manual inspections, machine inspections, and intelligent inspections [1, 2]. Among these, manual inspections are time-consuming and labor-intensive, and it is difficult to ensure the accuracy and comprehensiveness of the inspections [3]. Machine inspections utilize sensors and other devices, but issues such as delayed data analysis exist [4]. Intelligent inspections incorporate machine learning and other technologies into traditional inspections, effectively improving inspection efficiency and accuracy, but they are costly [5-7]. Therefore, power equipment inspection technology based on image recognition technology has emerged as a new option [8].

Power equipment inspection technology based on image recognition technology primarily involves collecting images or videos of power equipment for processing, and using machine learning and other technologies to assess and predict equipment status [9-11]. The specific steps include equipment data collection, image processing, feature extraction, model training, and equipment status assessment. Power equipment inspection technology based on image recognition has broad application prospects in fields such as automated production, coal mine

*tempmail0808@163.com
https://doi.org/10.65102/is2026663
Published: 30 April, 2026

safety, and urban power supply, particularly image recognition technology based on deep residual networks [12-15].

Deep residual networks (ResNet) are a method for training extremely deep networks. Unlike traditional convolutional neural networks (CNN), ResNet proposes a cross-layer residual learning approach, which avoids gradient vanishing or explosion issues through residual blocks and significantly alleviates the accuracy degradation caused by increasing network depth [16-19]. Traditional CNNs or fully connected networks inevitably suffer from information loss or degradation during information transmission, as well as issues like gradient vanishing or exploding gradients, leading to performance degradation when training deep CNNs [20-23]. To address this issue, an improved deep residual network is employed for image recognition, with simple shortcuts added to enhance the accuracy of image classification tasks and accelerate the training process, achieving greater precision in the identification of power equipment [24-26].

Reference [27] proposes a target detection network based on YOLOv4 and a two-stage fine-tuning method for power equipment image recognition. Experimental results demonstrate that the proposed method effectively enhances power equipment recognition capabilities. Literature [28] emphasizes the effective application of image recognition technology in online monitoring of power equipment and proposes a second template matching algorithm for rapid target image recognition. This algorithm is used to identify power equipment and detect abnormal states of power equipment, verifying that it can accurately locate and identify power equipment and detect equipment failures. Literature [29] addresses the limitations of traditional image recognition methods for power equipment by proposing a network based on an improved attention mechanism for image detection and recognition of power equipment. This method enhances the recognition and extraction of detailed target features, thereby improving recognition accuracy. Literature [30] develops an improved infrared defect recognition method for substation equipment based on traditional methods and conducts experiments using measured infrared images. The results validate the effectiveness and accuracy of this method. Literature [31] proposed a multi-label image recognition model for power equipment maintenance based on a multi-scale dynamic graph convolutional network, and compared and validated the experimental results using an electrical equipment defect dataset, revealing that the model exhibits outstanding performance. Literature [32] proposed an image recognition and processing model for electrical inspections, trained based on the latest advancements in neural networks, capable of object recognition, thereby enhancing inspection, operation, and maintenance efficiency while reducing costs. Reference [33] proposed a method for identifying and diagnosing power equipment faults based on convolutional neural networks. The effectiveness of this method was verified by comparing the error detection performance and image recognition accuracy within a time interval. Reference [34] addresses the growing complexity of power grid structures, which poses significant challenges for power personnel conducting inspections. To address issues such as low accuracy in traditional methods for fault identification, the study employs improved techniques like region-based fully convolutional networks, incorporates deep learning, and validates that this approach achieves higher detection accuracy. Literature [35] Based on machine learning technology, combined with machine learning operations under the concept of neural networks, this study explored the classification and recognition of image data of transformers, circuit breakers, and isolating switches in substations, revealing the effectiveness of the aforementioned methods. Literature [36] analyzed adversarial sample generation techniques for power images and developed an artificial intelligence model resilience assessment system to evaluate resilience against various hacker and organizational attacks. This system promotes the safe, effective, reliable, and stable

operation of artificial intelligence models in power scenarios. The aforementioned studies proposed power equipment image recognition technologies based on YOLOv4, convolutional neural networks, and machine learning, effectively improving the detection efficiency of power equipment and ensuring the safe operation of the power grid.

Addressing the issues of low efficiency and inconsistent diagnostic standards associated with traditional infrared inspection methods that rely on manual labor, this paper innovatively proposes an adaptive deep residual diagnosis model framework tailored for infrared inspection scenarios in substations. First, a finely annotated multi-state classification dataset of thermal defects in substation equipment was constructed. Subsequently, a lightweight residual network architecture was designed, employing convolution kernel decomposition technology and multi-scale feature fusion mechanisms to reduce model parameter size and enhance the model's ability to extract thermal defect features. A constraint-oriented Bayesian hyperparameter optimizer is developed to autonomously optimize key parameters. Through comparative experiments and thermal defect region instance identification experiments, the feasibility of the deep residual network-based thermal defect diagnosis model for substation equipment proposed in this paper is validated.

2 Thermal defect diagnosis model for substation equipment based on deep residual networks

When insulation defects occur in power transformation equipment, they typically manifest as an increase in equipment temperature. Based on the correlation between infrared images and temperature, a neural network is employed to input infrared images of different defect types into the model for supervised training, enabling the network to autonomously learn data features. Once the model is trained, inputting the infrared images of the equipment to be inspected directly allows for intelligent diagnosis of the equipment type and thermal status. Therefore, the construction of the diagnostic model in this paper primarily involves the establishment of a database and the design of the model architecture.

2.1 Construction of infrared image dataset for substation equipment

Deep learning-based image recognition methods require reliable data support. To meet the needs of this study, we collected and organized a total of 11,000 infrared images of 15 types of common substation equipment in normal operating conditions. The data was collected from precise temperature measurement historical data of over 30 substations ranging from 110kV to 500kV across different seasons. The shooting environments and personnel varied significantly, providing robust raw data support for the development of robust image analysis and recognition algorithms.

This paper focuses on typical equipment defect types, selecting a total of 600 defect samples with complete overall equipment images, clear heat sources, and complete temperature and visible light data. The typical defect sample dataset was then categorized and organized based on equipment type and defect type.

All infrared image data was processed and annotated according to the following steps:

1) Remove redundant information. Using the infrared image processing software Flir Tool, a batch processing program was developed to remove text and symbols from the original infrared images, preventing invalid information from interfering with image analysis and recognition results;

2) Annotate device type. Each infrared image was classified and annotated based on the type of target device it captured, excluding images where the target device was unclear or not of the

type of interest in this study. This annotation information was used for analyzing image data by device type and training and testing device classification algorithms;

3) Annotate defect severity. For all images in the typical defect sample dataset, locate the corresponding inspection report for the device and annotate them using the classification categories “general defect,” “severe defect,” and “critical defect.” For samples in the normal device dataset, annotate them as “no defect”;

4) Extract temperature maps. For all samples in the typical defect sample dataset, use infrared analysis software to export their temperature maps. For samples in the normal device dataset, randomly select 200 samples from each device type to extract temperature maps;

5) Label device locations. For all images where devices of the type of interest in this paper can be clearly distinguished, label them with rectangular boxes;

6) Data set cleaning. Re-check the annotation results of all data and remove image data with annotation errors, unclear main device types, severe defocusing, or questionable defect diagnosis results;

7) Export visible light images. The original image files saved by the thermal imager contain corresponding visible light images. Export the visible light images and copy their corresponding annotation results.

Image preprocessing methods:

Convolutional neural network models typically require fixed input image dimensions, generally set as square arrays with equal width and height. Images that do not meet the size requirements are typically scaled directly to the required dimensions. Handheld thermal imagers used for infrared temperature measurement commonly use image resolutions of 4:3 or 5:4. Direct scaling to a square array alters the original aspect ratio. The shooting angles of infrared images of power transformation equipment are highly random, and the direction in which the equipment is stretched during scaling is also random. Therefore, altering the aspect ratio can cause unpredictable edge deformation and alter the original edge information.

Changes in edge information significantly impact the resolution of substation equipment. Different types of substation equipment share many similar structures, such as insulators, voltage equalization rings, and connectors. The key basis for distinguishing these structurally similar devices visually is the differing size ratios of their components. Therefore, random angle-based aspect ratio transformations pose significant challenges for device differentiation. Maintaining the original aspect ratio of the image preserves critical edge information.

Deep learning heavily relies on data volume, so the training process of convolutional neural networks typically involves a series of preprocessing steps to enhance the dataset, i.e., generating new images from existing ones to improve the model's generalization ability across different processed images. One category of dataset enhancement methods involves adjusting the color distribution of images, such as contrast adjustment, normalization, and whitening. The purpose of these operations is to make the data distribution more uniform, enabling the recognition model to be more sensitive to different input ranges and adapt to visible light images under varying lighting conditions.

In summary, for infrared images of power transformation equipment, dataset augmentation methods such as rotation, cropping, and proportional scaling can be used to assist the model training process, while methods that cause aspect ratio changes or nonlinear grayscale adjustments should be avoided.

2.2 Infrared thermal imaging diagnosis model based on deep residual networks

Residual networks [37], as classic convolutional neural networks, have achieved good results in classification and recognition in various databases. However, due to the characteristics of

infrared images of power transformation equipment, such as complex backgrounds, similar shapes of some equipment, and different manifestations of defect states, if the ResNet network is directly used for thermal defect diagnosis of power transformation equipment in infrared images, the recognition accuracy is not high. Therefore, based on the characteristics of the infrared image dataset, this paper improves the residual network from three aspects: convolution kernel decomposition, multi-dimensional feature fusion, and convolution kernel self-adjustment, and designs the network structure of this paper.

The initial weight values of the convolution layer are set using the Xavier normal distribution initialization method to avoid the gradient vanishing problem caused by Gaussian distribution initialization, i.e.:

$$W_l = \text{random}N(\mu = 0, \sigma^2 = \frac{1}{p_{l-1}}) \tag{1}$$

In the equation: W_l is the initial weight parameter of layer l; p_{l-1} is the number of neurons in layer $l-1$; $N(\mu, \sigma^2)$ is the normal distribution function, μ is the mean of the random variable, and σ^2 is the variance of the random variable.

The output of each convolutional layer is normalized (BN) [38] and undergoes Relu activation. Normalization of the output accelerates network convergence, while the Relu activation function is used for feature mapping.

2.2.1 Network Structure Optimization Based on Convolution Decomposition

The imbalance in sample categories caused by the small number of defective images makes the model prone to overfitting during training, thereby reducing recognition accuracy. This paper uses convolution kernel decomposition technology to reduce the number of parameters, thereby improving the generalization performance of the network. The convolution kernel decomposition process is as follows:

$$(2n_c+1)\times(2n_c+1) \rightarrow \underbrace{(3\times3),\cdots,(3\times3)}_{n_c} \tag{2}$$

$$n_c \times n_c \rightarrow 1\times n_c + n_c \times 1 \tag{3}$$

In the equation: $n_c \times n_c$ represents the convolution kernel size; $\rightarrow$ denotes the direction of convolution kernel decomposition, with the convolution effect remaining unchanged before and after decomposition.

In the I-RB module structure, the larger 5×5 convolution kernel is first decomposed into two consecutive 3×3 convolution kernels. After convolution decomposition, the increase in network structure complexity is accompanied by an increase in training time. Therefore, only one of the 3×3 convolution kernels is asymmetrically decomposed into two consecutive convolution kernels of 3×1 and 1×3. The 1×1 convolution kernel is used to obtain feature dimension reduction, thereby accelerating the network training process.

2.2.2 Multi-scale convolution feature fusion optimization

The overheated areas caused by current-induced thermal defects in power transformation equipment account for a small proportion of the image. The characteristic information reflecting abnormal overheating in such equipment is primarily distributed in the shallow layers of the

network. After extraction and merging through multiple convolutional layers, this information is prone to loss, affecting the network's diagnostic results. To address this issue, a three-level hierarchical fast connection method is employed to highlight overheating feature information in the output high-level semantic features, thereby improving the model's diagnostic accuracy for overheating defects in power transformation equipment.

In the configuration of the fast connection method, the first level connects a single I-RB module to an RB module, the second level connects a single residual unit (where the I-RB module and RB module are integrated into a single residual unit), and the third level connects all residual units. The computational expression for each module's input is:

$$\begin{cases} x_{2i} = f(x_{2i-1} + F_1(x_{2i-1})), & i \in \{1, 2, \cdots, N/2\} \\ x_{2i+1} = f(x_{2i-1} + x_{2i} + F_2(x_{2i})), & i \in \{1, 2, \cdots, (N-2)/2\} \\ x_{N+1} = f(x_1 + x_N + x_{N-1} + F_2(x_N)), & N \in \{2, 4, \ldots, 2j\} \end{cases} \tag{4}$$

In the equation: x_i and x_{i+1} are the input and output of the i th module, respectively; F_1 and F_2 are the residual functions of the I-RB module and RB module, respectively; f is the Relu activation function; N is the total number of modules; j is a natural number.

In complex infrared image backgrounds, the contour distinction of power transformation equipment is not high. Increasing the network depth or the number of convolutional kernels can improve the network's ability to distinguish infrared images of power transformation equipment. Therefore, the network depth and the number of convolutional kernels are adjusted adaptively, with the network depth being increased sequentially and the number of convolutional kernels being adjusted according to the following principle:

$$k = 32 + 16a, a \in (0, 1, 2, 3, 4) \tag{5}$$

In the formula, k is the number of convolution kernels.

2.3 Model hyperparameter self-adjustment based on improved Bayesian methods

2.3.1 Bayesian Optimizer Principles and Implementation

1) Bayesian Optimizer Principle

Bayesian optimization [39] is an intelligent optimization method that can approximate the distribution of segmentation algorithm performance in thermal defect detection by constructing a proxy model (usually using a Gaussian process, GP). Combined with a collection function, this method can intelligently select the optimal evaluation point for segmentation algorithm parameters (such as thresholds, filter parameters, etc.), reducing the number of actual calculations and improving optimization efficiency. In thermal defect detection, Bayesian optimization effectively addresses the issue of insufficient segmentation accuracy in complex environments (e.g., varying lighting conditions, noise interference) through iterative updates of the model and parameters, ultimately significantly enhancing the adaptability and accuracy of the detection algorithm.

2) Bayesian optimization process

In thermal defect detection, the performance of the algorithm often depends on the precise adjustment of multiple parameters, such as the threshold selection of the Otsu threshold segmentation algorithm, the dual threshold settings of the Canny edge detection, and the kernel

size and standard deviation of the Gaussian filter. These parameters have a significant impact on the accuracy, robustness, and real-time performance of the detection results. However, the complex and variable real-world application environment of thermal defect detection (e.g., uneven lighting, noise interference, and dynamic backgrounds) makes traditional manual parameter tuning methods time-consuming and limited in effectiveness.

Parameter tuning methods based on the Bayesian optimization framework can transform the hyperparameter selection problem into an optimization process for the objective function, using segmentation performance metrics (e.g., F1 score or area under the precision-recall curve) as the optimization target. This optimization process includes the following key steps:

① Initialization stage: Set the initial number of sampling points n_0, determine the maximum number of evaluations N_1, and construct the initial training set D_0;

②Probabilistic proxy model construction: Use Gaussian process regression $G(K,d;\theta)$ to establish a probabilistic representation of the objective function, where: K is the kernel function, d is the input dimension, and θ is the model hyperparameter;

③Multi-objective sampling strategy: Design a robust sampling function as shown in Equation (6):

$$\alpha(K,d/D)=\mu(K,d)+\kappa\sigma(K,d) \tag{6}$$

where K is the exploration-exploitation balance coefficient.

④ The iterative optimization process is as shown in Equation (7):

$$\begin{gathered} while\,|D|<N_1 d_0 \\ (K^*,d^*)=\arg\max\alpha(K,d\,|\,D) \\ U^*=evaluate(K^*,d^*) \\ D\leftarrow D\cup\{(K^*,d^*),U^*\} \\ endwhile \end{gathered} \tag{7}$$

The parameter optimization flowchart is shown in Figure 1:

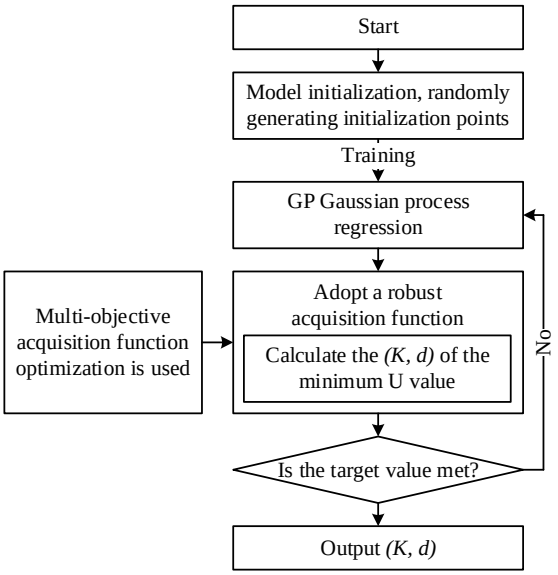


Figure 1: Flowchart of Bayesian optimization parameters

2.3.2 Coupling Constraint Settings

This paper imposes coupled constraints on the Bayesian optimization algorithm from two aspects: the size and accuracy of the diagnostic model. This biases the algorithm toward selecting hyperparameter combinations with higher expected accuracy and smaller network structures, making it suitable for thermal defect inspection. The constraints are as follows:

1) Constraint 1: The higher the validation set accuracy, the higher the feasibility evaluation value. After adding the constraint conditions, the acquisition function adjusts the feasible domain of the hyperparameter combinations based on the feasibility evaluation, thereby accelerating the convergence process of hyperparameter optimization. The improved acquisition function is such that when the accuracy is less than 80%, the optimization algorithm does not use that set of hyperparameters for updates.

2) Constraint 2: The fewer the number of network layers, the higher the feasibility evaluation value.

3) Constraint 3: The fewer the number of convolutional kernel channels, the higher the feasibility evaluation value.

$$\theta_{i+1} = \arg\max_{\theta \in R^*} E_D(F(\theta_{best}) - F(X_{i+1,j}))\prod_{t=1}^{T} C_t(X_{i+1,j}) \tag{8}$$

In the equation: θ_{i+1} is the hyperparameter combination for the $i+1$ th evaluation of model performance; R^* is the hyperparameter combination space; E_D is the function for calculating the incremental expectation based on the current dataset; F is the predictive evaluation function of the probabilistic proxy model for hyperparameter combinations; C_t represents the evaluation function for the t th constraint; T is the total number of constraints; θ_{best} is the optimal hyperparameter combination so far; $X_{i+1,j}$ is the j th candidate point in the $i+1$ st evaluation.

2.3.3 Bayesian algorithm optimization process based on coupling constraints

The Bayesian optimization process under coupling constraints is shown in Figure 2 and can be roughly divided into the following four steps:

1) Determine the parameter combinations to be optimized, set the hyperparameter optimization space, constraint conditions, and maximum optimization time.

2) Perform m random hyperparameter optimization training runs to obtain the model validation set accuracy and constraint evaluation values.

3) Update the non-parametric Gaussian process model and constraint evaluation model, and use the expected increment maximization function to select the next most “promising” hyperparameter combination.

4) After reaching the maximum optimization time, output the set of hyperparameter combinations.

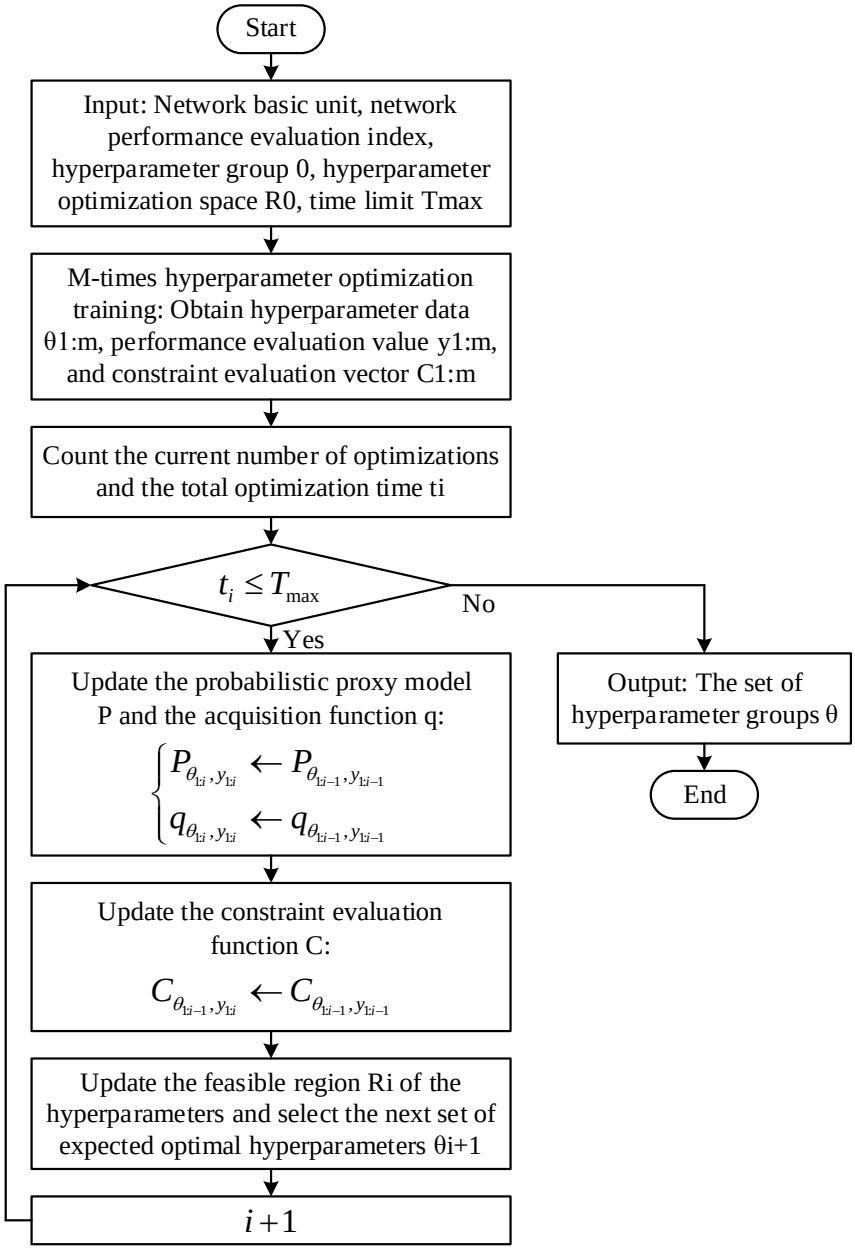


Figure 2: Flowchart of Bayesian optimization Algorithm based on coupling constraints

3 Thermal defect diagnosis test results and analysis

3.1 Evaluation Indicators

To evaluate the effectiveness of thermal defect detection, the average intersection-over-union ratio (MIoU) and precision rate are used to assess the performance of target recognition in infrared images. MIoU is one of the metrics for evaluating semantic recognition performance. It reflects the degree of overlap between the recognition results and the true values by calculating the ratio of the intersection and union (IoU) of two sets. In image semantic recognition, these two sets are the true values (labels) and the predicted values. The value of MIoU ranges from [0, 1], with higher values indicating better recognition performance. The definition of MIoU is as follows:

$$MIoU = \frac{1}{k+1}\sum_{i=0}^{k}\frac{p_{ii}}{\sum_{j=0}^{k}p_{ij}+\sum_{j=0}^{k}p_{ji}-p_{ii}} \tag{9}$$

In the formula: p_{ij} and p_{ji} represent the total number of pixels with predicted results of i and actual results of j, and the total number of pixels with predicted results of j and actual results of i, respectively, while p_{ii} represents the total number of pixels with predicted results of i and actual results of i.

Precision represents the category pixel accuracy rate of semantic recognition, which is the probability that a sample predicted as positive is actually a positive sample. Precision is defined as:

$$Precision = \frac{TP}{TP+FP} \tag{10}$$

In the formula: TP and FP represent true positives and false positives, respectively, indicating the proportion of examples classified as positive that are actually positive.

3.2 Network Training

The model training platform operates on the Ubuntu 18.04 operating system and utilizes an improved ResNet network built using the TensorFlow deep learning framework. The hardware processor is an Intel(R) Core (TM) i5-9300H CPU @2-4.0GHz (2400MHz), and the graphics card model is an NVIDIA GeForce GTX 1660 Ti. This paper uses the Res-UNet series of networks as a comparison network for thermal defect identification performance. To test the performance of different networks in identifying electrical equipment in complex environments, the network in this paper is compared with the UNet network and the Deeplabv3+ network. Relevant literature indicates that the Deeplabv3+ network outperforms the SegNet and FCN series networks in identifying current transformers in infrared images. Therefore, this paper chooses to compare with the Deeplabv3+.

The UNet, Deeplabv3+, and this paper's networks were all built on the TensorFlow learning framework. A total of 2,000 labeled infrared images were used as training samples, with the training set and test set randomly allocated at a ratio of 4:1. The sample images included current transformers and circuit breakers. The network proposed in this paper, UNet, and Deeplabv3+ all use cross-entropy as the loss function, defined as follows:

$$E = \sum_{c}^{C}\omega_c \log_{10}(p_c(z_i)) \tag{11}$$

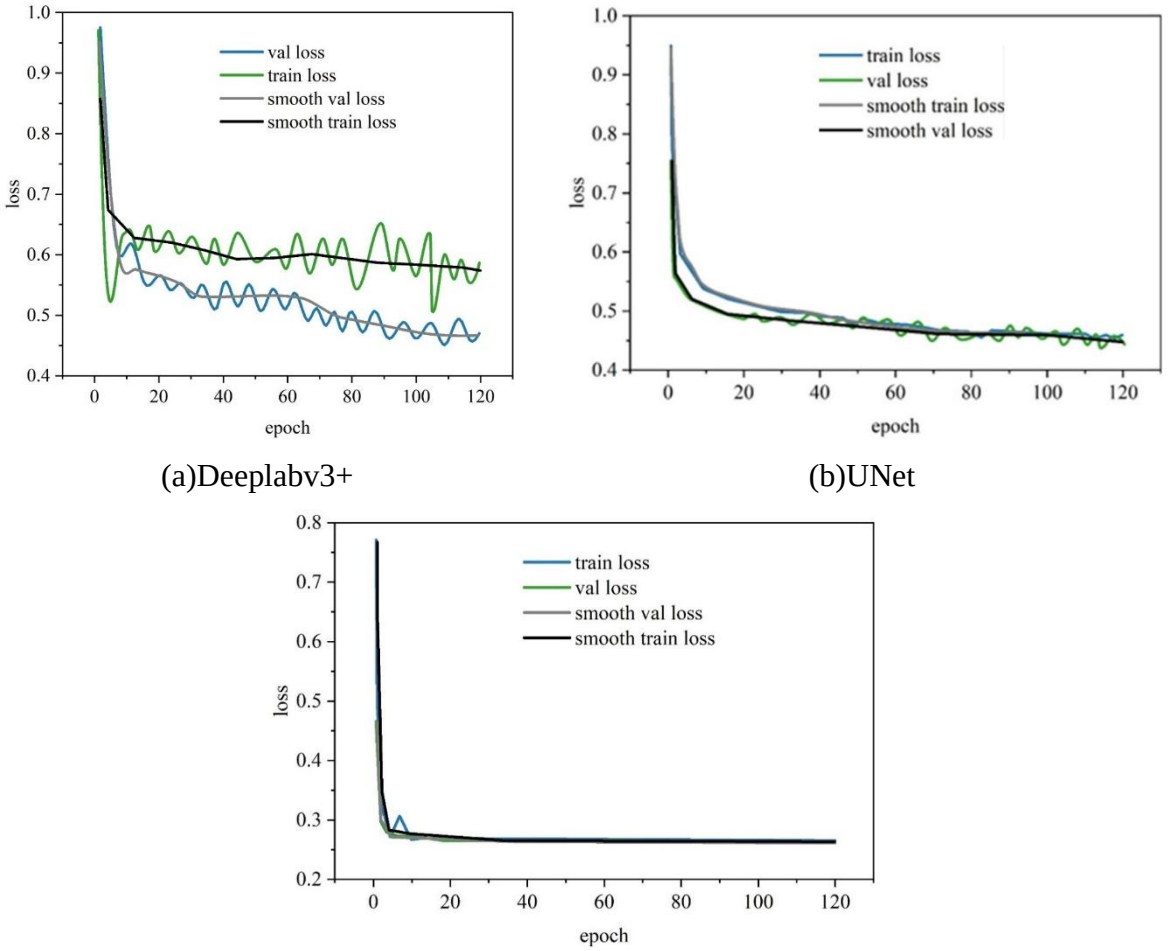
In the formula: ω_c is the loss weight of category c; $p_c(z_i)$ is the probability that pixel z_i belongs to the true category c.

3.3 Thermal defect diagnosis results

3.3.1 Comparison of recognition performance between different methods

The loss function curves obtained by training the three methods—the proposed model, UNet, and Deeplabv3+—are shown in Figure 3. (a), (b), and (c) represent the loss function

comparisons of the training processes for Deeplabv3+, UNet, and the proposed model, respectively. The Deeplabv3+ model did not converge by epoch 120, while the model proposed in this paper achieved convergence by epoch 12, demonstrating faster convergence and a loss value closer to 0 at convergence.



(a)Deeplabv3+ (b)UNet

(c)This model

Figure 3: Loss function curve

The same samples are used to train different network models to obtain network parameters. Test samples are then input to obtain recognition prediction results, and the performance of infrared image target recognition is evaluated using MIoU and Precision. Cases 1–4 involve randomly selecting some samples from the test set to test recognition performance. Cases 1–4 all use infrared images captured at the substation selected in this paper. Table 1 shows the MIoU values corresponding to the recognition results for Cases 1–4. The recognition targets for Scenarios 1 and 2 are thermal defects in current transformers, while the recognition targets for Scenarios 3 and 4 are thermal defects in circuit breakers. As shown in the results of Table 1, the recognition performance of the UNet network is superior to that of the Deeplabv3+ network, and the network selected in this paper outperforms the other four networks. In Case 1, the background of the current transformer is simple, and the target recognition rate of the network in this paper is 0.9312. In Case 2, the current transformer is in a complex environment, with other electrical equipment present in addition to the current transformer. The network in this paper can accurately identify the target in a complex background, achieving a recognition accuracy of 0.8987. Scenarios 3 and 4 both include multiple circuit breakers, with some circuit

breakers partially obstructed or interfered with by their background, and the textures of the circuit breakers are unclear. The network achieves a thermal defect recognition accuracy rate above 0.9 for both scenarios, successfully identifying thermal defects in the circuit breakers.

Table 2 presents the statistical results of the recognition accuracy of different networks on a test dataset of 400 images. The test samples are categorized into three classes: current transformers, circuit breakers, and background. MIoU and accuracy are used to measure the accuracy of identifying thermal defects in the three target categories across different samples. The identification results are shown in Table 2. As indicated by the data in Table 2, the UNet network achieves more accurate results than the Deeplabv3+ network. Compared to the other four networks, the proposed network demonstrates superior performance in identifying thermal defects in both types of electrical equipment. However, the identification performance of the Res34-UNet and Res50-UNet networks is worse than that of the UNet segmentation. It can be seen that the Res-UNet network, which uses the ResNet network as the code part to extract features and construct the UNet network, can indeed improve the accuracy of target identification. However, deep networks (such as Res34-UNet and Res50-UNet) suffer from overfitting due to the small number of training samples, which leads to a decrease in identification accuracy.

Table 1: The MIOU values based on different methods

Different situations	Deeplabv3+	UNet	This network	Res34-UNet	Res50-UNet
1	0.7890	0.8208	0.9312	0.7795	0.6212
2	0.7765	0.7875	0.8987	0.8182	0.6635
3	0.7911	0.8302	0.9015	0.7305	0.6328
4	0.7882	0.8263	0.9052	0.7162	0.6583

Table 2: The accuracy of the test dataset

Network	Segmentation object	Iou	Miou	Precision
Deeplabv3+	Current transformer	0.7912	0.8081	0.9010
	Circuit breaker	0.6784		0.8451
	Background	0.9546		0.9711
Unet	Current transformer	0.8023	0.8285	0.9158
	Circuit breaker	0.7178		0.8987
	Background	0.9654		0.9874
This network	Current transformer	0.8564	0.8953	0.9457
	Circuit breaker	0.8597		0.9345
	Background	0.9698		0.9948
Res34-unet	Current transformer	0.6302	0.7135	0.7154
	Circuit breaker	0.6054		0.7358
	Background	0.9048		0.9874
Res50-unet	Current transformer	0.4784	0.5918	0.5178
	Circuit breaker	0.4215		0.3785
	Background	0.8754		0.9688

3.3.2 Visualization Analysis of Network Characteristics

Figures 4 and 5 show the visualization of thermal defect detection for the thermal defect diagnosis model of substation equipment using Bayesian optimization and the thermal defect diagnosis model of substation equipment without Bayesian optimization, respectively. There are seven types of thermal defects in the figure, namely, abnormal overheating of bushings, abnormal temperature distribution of bushings, normal bushings, abnormal overheating of current transformers, normal current transformers, abnormal overheating of disconnectors, and

normal disconnectors. The method proposed in this paper can obtain clear classification boundaries for various types of thermal defect data, achieving high thermal defect diagnosis performance. However, as shown in Figure 5, in the thermal defect diagnosis model for substation equipment without Bayesian optimization, due to the suboptimal combination of hyperparameters, the boundaries between the fifth and sixth types of faults overlap, resulting in low accuracy of thermal defect detection.

A Bayesian optimization algorithm is constructed with thermal defect detection accuracy as the objective function to optimize the hyperparameters of the network in the thermal defect detection model. This method can effectively detect and classify thermal defects under various thermal defect patterns, aiding in the accurate analysis of thermal defects in power equipment and providing stronger technical support for the safe operation of power equipment.

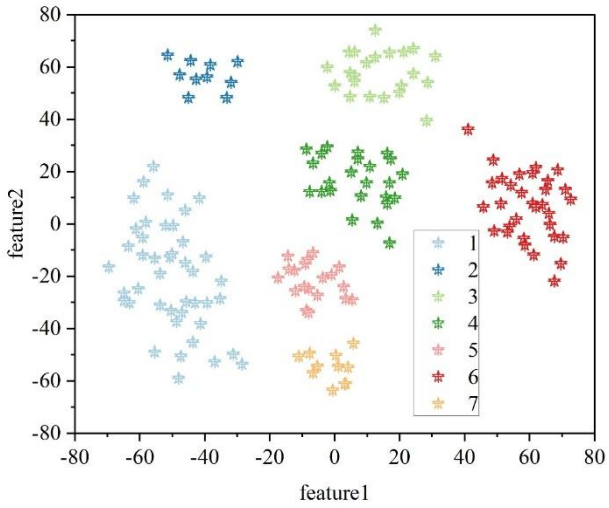


Figure 4: Visual expression of this article

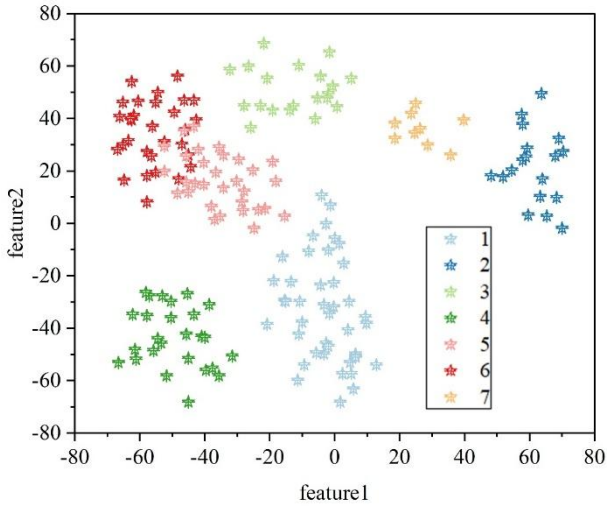


Figure 5: The visual expression of beesus optimization was not adopted

3.3.3 Overheating area identification results

The three temperature thresholds determined based on the relative temperature difference criterion for substation equipment are 35%, 80%, and 95%. If the relative temperature difference is less than 35%, the equipment is considered to be operating normally. If the relative temperature difference is greater than 35%, it is considered that there is an overheating defect.

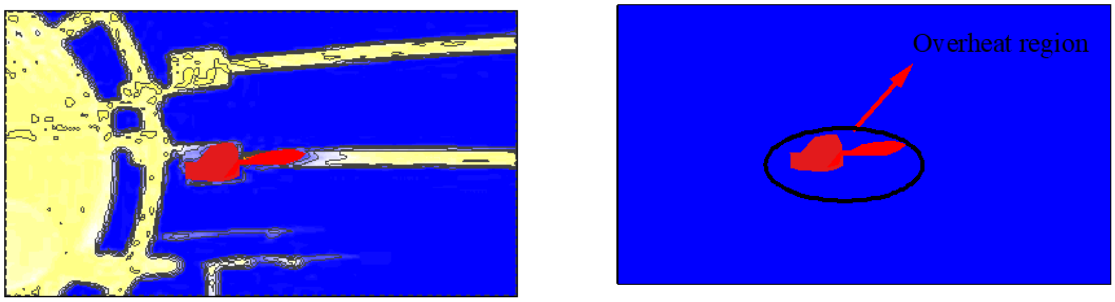
According to “DL/T 664-2008 Specifications for Infrared Diagnosis of Live Equipment,” defects can be classified into general defects, serious defects, and emergency defects. General defects refer to equipment that is overheating, but to a minor degree, and with a small temperature difference and temperature field gradient. Such defects generally do not cause accidents and can be eliminated during power outages for maintenance. Severe defects refer to cases where the equipment's overheating phenomenon is already severe, and both the temperature difference and temperature field gradient are significant. Such defects should be addressed promptly; otherwise, they may escalate into critical defects. Critical defects refer to cases where the equipment's overheating phenomenon has reached a critical state, requiring immediate power shutdown for resolution.

The classification of defects at various levels is shown in Table 3.

Table 3: Classification of defects at all levels

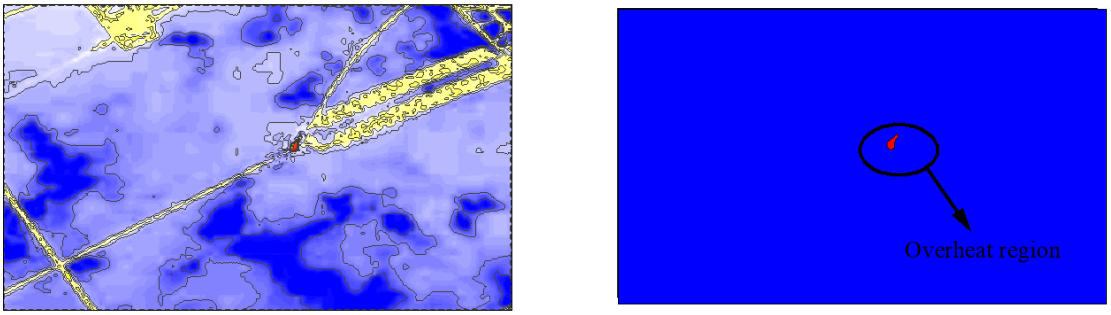
	Normal	General defect	Serious defect	Emergency defect
Relative temperature difference	X<35%	35%≤X<80%	80%≤X<95%	X>95%

This paper uses Matlab R2015a as the programming tool and selects 340×240-sized 500kV inductor bushing connectors, line connectors, and 220kV transformer boxes as the experimental objects. The fault diagnosis is performed using the thermal defect diagnosis model for power transformation equipment based on deep residual networks developed in this paper. Figures 6, 7, and 8 show the results of overheated region extraction along the bolts for the 500kV inductive bushing, line joint, and 220kV transformer box, respectively. Figures (a) and (b) show the original images and the overheated region extraction results, respectively.



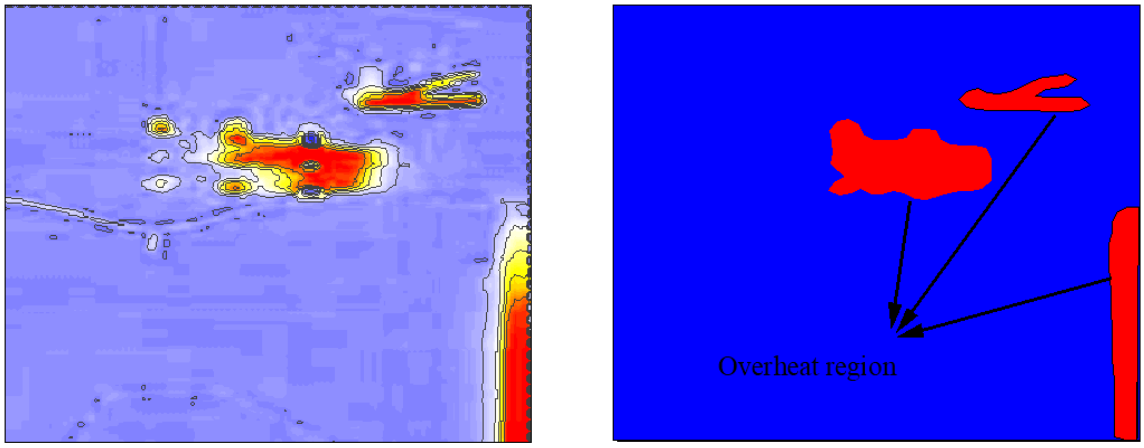
(a)Primary image (b)Overheat region

Figure 6: Inductance outlet connection



(a)Primary image (b)Overheat region

Figure 7: Line joint



(a)Primary image (b)Overheat region

Figure 8: Transformer box along bolt

Using the fault diagnosis algorithm described in this paper, calculate the relative temperature difference and determine whether the power equipment in Figures 6, 7, and 8 has overheating faults based on Table 3. Table 4 provides the relative temperature difference values for the equipment.

The known equipment faults are: loose joint crimping causing heating in the 500 kV inductor, poor contact at the line joint causing heating, and leakage magnetic flux from the 220 kV transformer tank along the bolts causing bolt heating. By comparing with the data in Table 3, it can be seen that the power equipment in Figures 6 and 7 have both experienced overheating faults, and after diagnosis, they are both classified as minor defects. The relative temperature difference of the power equipment fault in Figure 8 is 84.55%, which is greater than 80%, and is diagnosed as a severe defect. All should be addressed during shutdown maintenance. The diagnostic results are consistent with the actual fault conditions of the known equipment, and the algorithm's detection performance is good.

Table 4: Relative temperature difference

	Ambient temperature(℃)	Normal regional temperature(℃)	Heat of overheat(℃)	Relative temperature difference/%
Inductance outlet connection	25.2	28.4	32.1	53.62
Line joint	6.4	26.4	84.9	74.52
Transformer box	20.3	25.8	55.9	84.55

4 Conclusion

(1) This paper addresses the issue of limited sample sizes for thermal defect samples in power equipment by introducing convolutional decomposition to reduce model complexity and parameter counts. On the other hand, by utilizing multi-scale feature fusion technology, the model can capture thermal distribution information at different scales, thereby improving the accuracy of the thermal defect diagnosis model for power equipment based on deep residual networks in identifying thermal defect states of power equipment. In simple background scenarios, the thermal defect identification accuracy of this model reaches 0.9312.

(2) The improved Bayesian model hyperparameter auto-adjustment algorithm introduces coupled constraint settings, enabling the model to converge more efficiently to the optimal hyperparameter region. This allows the model to achieve optimal performance on a limited dataset, accurately classifying seven different types of thermal defects, and is suitable for intelligent recognition tasks involving thermal defect inspection images of electrical equipment.

(3) The results of thermal defect region instance identification indicate that the thermal defect diagnosis model can diagnose thermal defects in power equipment caused by different reasons. For example, the relative temperature difference caused by overheating of the bolts along the transformer tank is 84.55%, which is a severe defect. This result is consistent with the actual fault situation. The method provides an effective technical means for power equipment condition monitoring and has important engineering application value.

About the Authors

Xiaolong Chen, male, 2011- 2017, South China University of Technology, majoring in Electrical Engineering and Automation, Ph.D. During the Ph.D. program, he researched on the impact of DC system control on AC systems. 2006 -2009, Guangdong University of Technology, majoring in Power System and Automation, he focused on AI-based control for hybrid AC/DC transmission systems. Since 2009, working at China Southern Power Grid, a senior engineer, with 16 years experience at substation operation, he is a senior technical lead and in charges of smart operations and maintenance upgrade projects and management. As the first author, he has published 6 journal papers, 2 of which are indexed by EI.

Chao Sun, male, 2010 - 2013, Zhejiang University, majoring in Power Electronics and Power Drives, Master's Degree. 2014 - 2018, Dalian University of Technology, majoring in Electrical Engineering and Automation, Bachelor's degree. Since 2013, working at Guangzhou Power Supply Bureau, China Southern Power Grid, he has participated in the development of multiple substation intelligent operation and maintenance systems and the reconstruction and construction projects of station-side facilities. These efforts have effectively improved the quality and efficiency of substation equipment operation and maintenance. Responsible for organizing and implementing projects such as the construction and application of intelligent inspection robot systems, the intelligent transformation of remote patrol and operation in substations, and the construction of integrated networking inspection systems for intelligent operation and maintenance of substations. He has been awarded the First Prize for Excellent Papers at the National Power Grid Intelligent Operation and Maintenance Conference and the Second Prize for Innovation Achievements in Equipment Management of the National Electric Power Industry.

Xuming Ni, male, 2008-2011, South China University of Technology, majoring in electrical engineering, master’s degree. His primary research focused on power sytem protection andcontrol. 2004-2008, South China Agricultural University, majoring in Electrical Engineering and Automation, Bachelor's degree. He focused on power grid load forecasting. As a senior engineer, he is the senior manager of substation operations. With 14 years of specialization in technical management of substation operations, he is engaging in research on the intelligent operations and maintenance technologies for substations. As the first author, he has published 5 journal papers, 2 of core power engineering journals, and participated in development and construction of China Southern Power Grid’s pioneering Substation Operational Support System.

Xiaoyuan Jia, male, 2008 - 2011, Xi’an Jiaotong University, majoring in Electrical Engineering and Automation, Master's degree. During the master's program, he focused on photovoltaic system research and proposed a novel planning method for photovoltaic power

plants. 2004 - 2008, Xi'an Jiaotong University, majoring in Electrical Engineering and Automation, Bachelor's degree, with outstanding academic performance and awarded multiple university-level scholarships. Since 2011, working at Guangzhou Power Supply Bureau as a substation operation manager, he has participated in the development of a new intelligent operation and maintenance strategy, which has been applied in several substations and effectively enhanced O&M quality and efficiency. As the first author, he has published one EI-indexed paper.

Xing Wang, female, 2007 - 2009, Wuhan University, majoring in Science in Power System and Automation, Master's degree. During the master's period, the main research focused on online monitoring of power systems, and an ultrasonic partial discharge monitoring method for transformers was proposed. 2003 - 2007, Wuhan University, majoring in Engineering in Electrical Engineering, Bachelor's degree. Since 2009, she is working in Substation Management Division I of Guangzhou Power Supply Bureau, Guangdong Power Grid, engaged in relay protection, substation production technology management and other related work.She has published 5 Chinese papers, with in-depth research in relay protection device calibration, secondary acceptance of circuits, fault analysis, etc. She has obtained 42 innovation awards from 2013 to 2023, and published 2 invention patents, 7 utility model patents, and 3 software copyrights.

Jian Pang, male, 2018 - 2021, North China Electric Power University, majoring in Electronic Science and Technology, Master's degree. His direction during the Master's degree program was the advanced transmission and substation technology. 2014 - 2018, North China Electric Power University, majoring in Smart Grid Information Engineering, Bachelor's degree. Since 2021, employed by Guangzhou Power Supply Bureau, he is working in substation operation and participating in the construction of digital substation production systems. He has published 1 paper in a Chinese core journal, 1 paper in an EI-indexed conference and obtained 1 national authorized invention patent during university.

References

[1] Yang, L., Fan, J., Liu, Y., Li, E., Peng, J., & Liang, Z. (2020). A review on state-of-the-art power line inspection techniques. IEEE Transactions on Instrumentation and Measurement, 69(12), 9350-9365.

[2] Siddiqui, Z. A., Park, U., Lee, S. W., Jung, N. J., Choi, M., Lim, C., & Seo, J. H. (2018). Robust powerline equipment inspection system based on a convolutional neural network. Sensors, 18(11), 3837.

[3] Sun, H., Lu, S., & Feng, T. (2023, October). Electric power inspection equipment: overview, challenges, and applications. In 12th International Conference on Renewable Power Generation (RPG 2023) (Vol. 2023, pp. 1179-1184). IET.

[4] Zhang, T., & Dai, J. (2021). Electric power intelligent inspection robot: a review. In Journal of Physics: Conference Series (Vol. 1750, No. 1, p. 012023). IOP Publishing.

[5] Kruglova, T. N. (2015). Intelligent diagnosis of the electrical equipment technical condition. Procedia Engineering, 129, 219-224.

[6] Eltyshev, D. K., & Kostygov, A. M. (2019). Intelligent diagnostic control and management of the condition of electrotechnical equipment. Russian Electrical Engineering, 90(11), 741-746.

[7] Zheng, H., Ping, Y., Cui, Y., & Li, J. (2022). Intelligent diagnosis method of power equipment faults based on single-stage infrared image target detection. IEEJ Transactions on Electrical and Electronic Engineering, 17(12), 1706-1716.

[8] Wang, H., & Meng, F. (2019). Research on power equipment recognition method based on image processing. EURASIP Journal on Image and Video Processing, 2019(1), 57.

[9] Haider, M., Doegar, A., & Verma, R. K. (2018, September). Fault identification in electrical equipment using thermal image processing. In 2018 International conference on computing, power and communication technologies (GUCON) (pp. 853-858). IEEE.

[10] Xu, J., Xu, Y., Zhu, Z., Hu, Z., Hu, S., & Li, Z. (2024, April). Research on the Construction and Application of Visual Inspection System for Power Tools Based on Image Recognition Technology. In 2024 IEEE 2nd International Conference on Control, Electronics and Computer Technology (ICCECT) (pp. 1089-1092). IEEE.

[11] Ye, Q. (2021). Intelligent inspection system of power equipment based on photoelectric sensor/AR technology. Journal of Nanoelectronics and Optoelectronics, 16(10), 1645-1656.

[12] Zhang, H., Wu, G., Ma, A., Ma, Y., & Zou, A. (2017). Application of infrared image recognition technology in electrical equipment. DEStech. Trans. Eng. Technol. Res., 38(3), 88-91.

[13] Li, J., & Wang, X. (2024, March). Deep learning-based image recognition for power line inspection. In International Conference on Computer Graphics, Artificial Intelligence, and Data Processing (ICCAID 2023) (Vol. 13105, pp. 146-157). SPIE.

[14] Wang, Y. (2023). Electrical Control Equipment Patrol Inspection Method Based on High Quality Image Recognition Technology. Traitement du Signal, 40(2).

[15] Zhang, Q., Chang, X., Meng, Z., & Li, Y. (2021). Equipment detection and recognition in electric power room based on faster R-CNN. Procedia Computer Science, 183, 324-330.

[16] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).

[17] He, L., Tiancheng, Z., Junbo, L., Lixin, J. I. A. O., Zhihao, X. U., & Xiaocui, Y. U. A. N. (2022). Deep Residual UNet Network-based Infrared Image Segmentation Method for Electrical Equipment. Infrared Technology, 44(12), 1351-1357.

[18] Alaeddine, H., & Jihene, M. (2021). Deep residual network in network. Computational intelligence and neuroscience, 2021(1), 6659083.

[19] Shafiq, M., & Gu, Z. (2022). Deep residual learning for image recognition: A survey. Applied sciences, 12(18), 8972.

[20] Hijazi, S., Kumar, R., & Rowen, C. (2015). Using convolutional neural networks for image recognition. Cadence Design Systems Inc.: San Jose, CA, USA, 9(1), 39.

[21] Li, Q., Cai, W., Wang, X., Zhou, Y., Feng, D. D., & Chen, M. (2014, December). Medical image classification with convolutional neural network. In 2014 13th international conference on control automation robotics & vision (ICARCV) (pp. 844-848). IEEE.

[22] Liu, Y. H. (2018, September). Feature extraction and image recognition with convolutional neural networks. In Journal of Physics: Conference Series (Vol. 1087, p. 062032). IOP Publishing.

[23] Ramprasath, M., Anand, M. V., & Hariharan, S. (2018). Image classification using convolutional neural networks. International Journal of Pure and Applied Mathematics, 119(17), 1307-1319.

[24] Tai, Y., Yang, J., & Liu, X. (2017). Image super-resolution via deep recursive residual network. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 3147-3155).

[25] Wang, F., Jiang, M., Qian, C., Yang, S., Li, C., Zhang, H., ... & Tang, X. (2017). Residual attention network for image classification. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 3156-3164).

[26] Zhou, F., Ma, Y., Ma, Y., & Pan, H. (2020, March). Infrared image fault identification of power equipment based on residual network. In Workshops of the International Conference on Advanced Information Networking and Applications (pp. 3-13). Cham: Springer International Publishing.

[27] Wu, J., Zeng, J., Zhou, Y., Zhang, Y., & Zhang, Y. (2023). Few-shot electrical equipment image recognition method based on an improved two-stage fine-tuning approach. The Journal of Engineering, 2023(9), e12313.

[28] Wu, G., Yu, M., Shi, W., Li, S., & Bao, J. (2020). Image recognition in online monitoring of power equipment. International Journal of Advanced Robotic Systems, 17(1), 1729881419900836.

[29] Lin, S. (2025). Power Equipment Image Recognition Method based on Feature Extraction and Deep Learning. IEEE Access, (99), 1-1.

[30] Luo, L., Ma, R., Li, Y., Yang, F., & Qiu, Z. (2021). Image recognition technology with its application in defect detection and diagnosis analysis of substation equipment. Scientific Programming, 2021(1), 2021344.

[31] Yan, Y., Han, Y., Qi, D., Lin, J., Yang, Z., & Jin, L. (2023). Multi-label image recognition for electric power equipment inspection based on multi-scale dynamic graph convolution network. Energy Reports, 9, 1928-1937.

[32] Xia, Y., Lu, J., Li, H., & Xu, H. (2018, October). A deep learning based image recognition and processing model for electric equipment inspection. In 2018 2nd IEEE Conference on Energy Internet and Energy System Integration (EI2) (pp. 1-6). IEEE.

[33] Xiaoxiong, L. (2025, February). Power Equipment Fault Image Recognition and Diagnosis Based on Convolutional Neural Network. In 2025 4th International Conference on Sentiment Analysis and Deep Learning (ICSADL) (pp. 1274-1279). IEEE.

[34] Lv, J., Zhang, H., Chen, X., Huang, J., & Zhou, H. (2024, January). An image recognition method for intelligent inspection of power grid equipment. In Proceedings of the 2024 International Conference on Power Electronics and Artificial Intelligence (pp. 463-467).

[35] Li, Z., & Su, H. (2021). Research on Image Recognition of Electrical Equipment based on Deconvolution Feature Extraction. In E3S Web of Conferences (Vol. 257, p. 01019). EDP Sciences.

[36] Chen, X., & Li, X. (2022, November). Research on robustness technology of power equipment image recognition algorithm model in confrontation scenario. In 2022 9th International Forum on Electrical Engineering and Automation (IFEEA) (pp. 1231-1235). IEEE.

[37] Minhaz Chowdhury,Shreoshi Roy Shrima & Md Shofiqul Islam. (2025). Comparative analysis of deep learning architectures for multi-class mineral classification: a study using EfficientNet and ResNet models. Earth Science Informatics,18(3),485-485.

[38] Go Eun Woo,Sang Bo Park,Gi Tae Park,Muhammad Junaid & Hyung Won Kim. (2025). Low-Complexity Hardware Architecture for Batch Normalization of CNN Training Accelerator. Computers, Materials & Continua,84(2),3241-3257.

[39] Juan Guan & Yanhua Wang. (2025). An improved high-dimensional Bayesian optimization algorithm. Applied Intelligence,55(13),915-915.