Thermal Anomaly Area Intelligent Recognition Method Based on Infrared Inspection Image of Electrical Equipment

Laiqi Wang [1], Jianhua Ou [1], Jianguo Wang [1*], Cong Hu [1], Yadong Fan [1], Fuhua Xie [1]

[1] School of Electrical Engineering and Automation, Wuhan University, Wuhan, China
[*] wjg@whu.edu.cn

Abstract: The thermal anomaly area of electrical equipment in the substation is often hidden due to its small thermal area and multiple anomalies overlaid. Accurately identify the thermal area is demanded on the condition detection of electrical equipment, where the anomaly points of electrical equipment in infrared images are generally small and of low resolution. We propose an improved YOLOv4 algorithm for infrared image anomaly area identification, which can detect the thermal generation phenomenon of electrical equipment. We add a new target detection branch to the shallow feature map of 104 × 104, which can better extract small target semantic information. The training process is enhanced with cosine annealing and mosaic data enhancement. We establish a total of 719 infrared images of five types of thermal anomalies electrical equipment to test our network. The accuracy of our model reach to as high as 96.78%, with a detection speed of 17 f/s and an AP@0.5 of 94.23%. Compared with SSD, YOLOv4 and Faster RCNN, the algorithm in this paper obtains the highest AP@0.5 with 94.23%, which is the best performance compared with the original YOLOv4 model in accuracy. The model is robust to noise and luminance disturbances, and still provides good recognition under disturbances.

***Index Terms*—Infrared image, Target detection branch, robustness test, Electrical equipment fault identification.**

1. INTRODUCTION

Electrical equipment undertakes important tasks in the substation , where it ensures the stable operation of substation and power system. Thermal faults may occur when the electrical equipment is in long-term operation with heavy load[1][2]. The thermal area of electrical equipment in the substation is often hidden [3], due to its area is often very small, and multiple devices may occur at the same time with a variety of faults. To accurately and completely identify the thermal area is very difficult [4][5]. This demand rises more complex requirements on the condition detection of electrical equipment.

When an electrical device fails in heavy load, it is usually accompanied by overheating. Infrared images can be taken by installing infrared camera equipment, where it can reflect the infrared image of overheating [6]. By detecting the hot spot on the infrared image, we can quickly find whether the electrical equipment is in the state of thermal anomalies.

There are researchers using machine learning methods to study infrared images of electrical equipment. Since machine learning requires a large amount of feature extraction work, the features in the anomalous regions of infrared images are difficult to be fully extracted, resulting in poor image recognition[7]. Yongbo Li et al [8] proposed a system for non-smooth running fault diagnosis of rotating machinery based on infrared thermography (IRT). Mohd Shawal Jadin et al [9] used the multilayer perceptron (MLP) artificial neural network and support vector machine (SVM) for classification.But none of their results were particularly satisfactory.

Compared with machine learning, deep learning can extract image features better. Deep learning networks can automatically train parameters based on a large number of images, saving a lot of human and material resources[10][11]. Deep learning networks are used for infrared image fault recognition in electrical equipment due to their excellent results in image recognition[12][13]. Haidong Shao [14] improved CNN by introducing random pools and leakage-corrected linear units to overcome the training problem in classical CNN. Faster RCNN has better detection effect and is used in electrical equipment fault detection. Some scholars have achieved over ninety percent mAP using Faster RCNN to detect electrical devices[15]. Lincong She[16] proposes a multi-scale residual network with convolutional filtering using convolutional kernels of different sizes to achieve insulator breakage identification. Faster RCNN has also achieved relatively good detection results on infrared images, Jianhua Ou *et al* of our team [17] improved VGG16 in Faster RCNN by adding anchors with aspect ratios of 1:3 and 3:1, and achieved better recognition results for infrared images.CNN and Faster RCNN have achieved good results in electrical equipment image recognition. However, CNN and Faster RCNN have more complex structure and training parameters, which lead to their slower computing speed.

YOLO algorithm that can better meet the needs of real-time monitoring of electrical equipment. Shenghui Wang *et al* [18] built a database of insulator images for training and testing, and used the YOLO algorithm to train the database for five days, achieving good recognition results. Lianqiao Li[19] designed a new feature enhancement module and proposed an infrared insulator image detection model based on an improved feature fusion single multibox detector. The model achieves an accuracy higher than 80% and improves detection speed.

However, the anomalies of electrical devices in infrared images are generally small and of low resolution. Current model for detection is less effective on this problem. To solve this problem we propose an improved YOLOv4 algorithm, where we add the detection branch to the 104×104 feature layer, since the shallow feature layer of the network retains more semantic information such as edges and textures. We collected a total of 719 infrared images of five types of

thermal faulty electrical devices to test our network. During the testing we used training strategies such as three-layer detection layer, four-layer detection layer, K-means clustering, cosine annealing and mosaic data enhancement. By comparing with other methods such as Faster RCNN and SSD, the results show that our model has high accuracy with very fast speed. After adding luminance and noise interference to the images, the model still has good recognition effect, which indicates that the model has strong robustness.

This paper is structured as follows, Section II introduces the improved YOLOv4 model and dataset structure. Section III describes the experimental test results and analysis of the results. Section IV concludes and outlooks the paper.

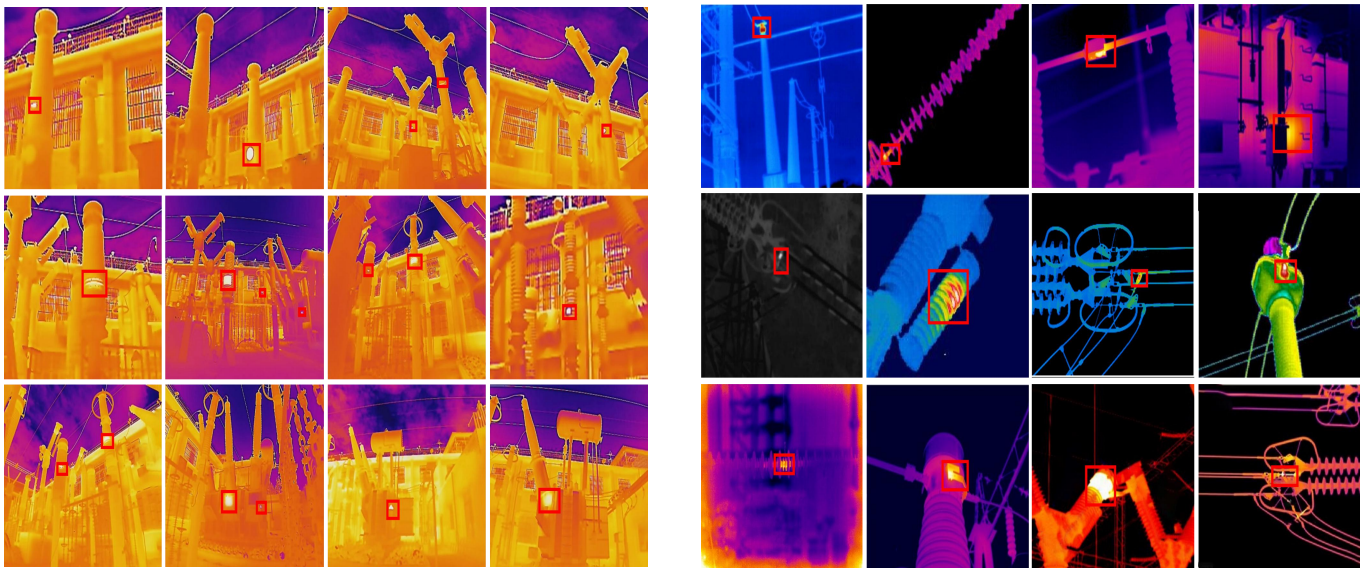
2. DATA AND METHODS

2.1. Infrared Image Dataset

The dataset in this paper contains a total of 719 initial images, most of which were taken from an electrical equipment thermal anomaly simulation experiment in a substation. We simulate different shapes of thermal anomalies on electrical devices by placing different shapes of thermal silicon wafers. We simulate thermal anomalies in various locations by placing thermal silicon wafers on different electrical devices. We also collect thermal anomalies of electrical equipment at different times to simulate realistic operating conditions. A total of 365 infrared images of electrical equipment with thermal faults were obtained after eliminating unclear heated areas and blurred images. Some images of the dataset are shown in Fig. 1(a). There are different kinds of thermal imagers used for infrared thermal fault detection of substation electrical equipment, and the heating fault conditions vary from substation to substation. The target detection model trained by a single thermal anomaly picture of substation electrical equipment obtained from an infrared thermal imager has poor robustness. 345 infrared images of electrical equipment faults were collected from the Internet to further enrich our training dataset. Some images of the dataset are shown in Fig. 1(b). The images collected in this paper contain thermal fault maps of insulator strings, transformers, current transformers, wire clamps, and other equipment in substations. A wide variety of thermal fault

TABLE I
DATA AUGMENTATION METHODS

Image processing methods		Parameters
Physical space change	Panning	X=[-30,+30], Y=[-30, +30]
	Rotation	(-20°,20°)
	Scaling	scale=(0.6,1.5)
	Horizontal Mirroring	With the right border as the axis
	Random cropping	percent=(0.1,0.3))
	Mis-cut	"x": (-20, 20), "y": (-20, 20)
	Gaussian blur	(sigma=(1,3))
	Up and down flip	With the upper border as the axis
	Sharpening	alpha=1, lightness=1
Color space transformation	Gray	Full image graying
	HIS	Hue ,Saturation,Intensity
	Lab	CIE Lab Color Space
	Luv	CIE Luv Color Space
	YCbCr	YCbCr color space



(a) (b)

Fig. 1. *Infrared image dataset of faulty electrical equipment. (a) Partial dataset for thermal fault simulation, all pictures are from Wuchang East 220kv substation. (b) Infrared images of electrical equipment failures on the Internet.*

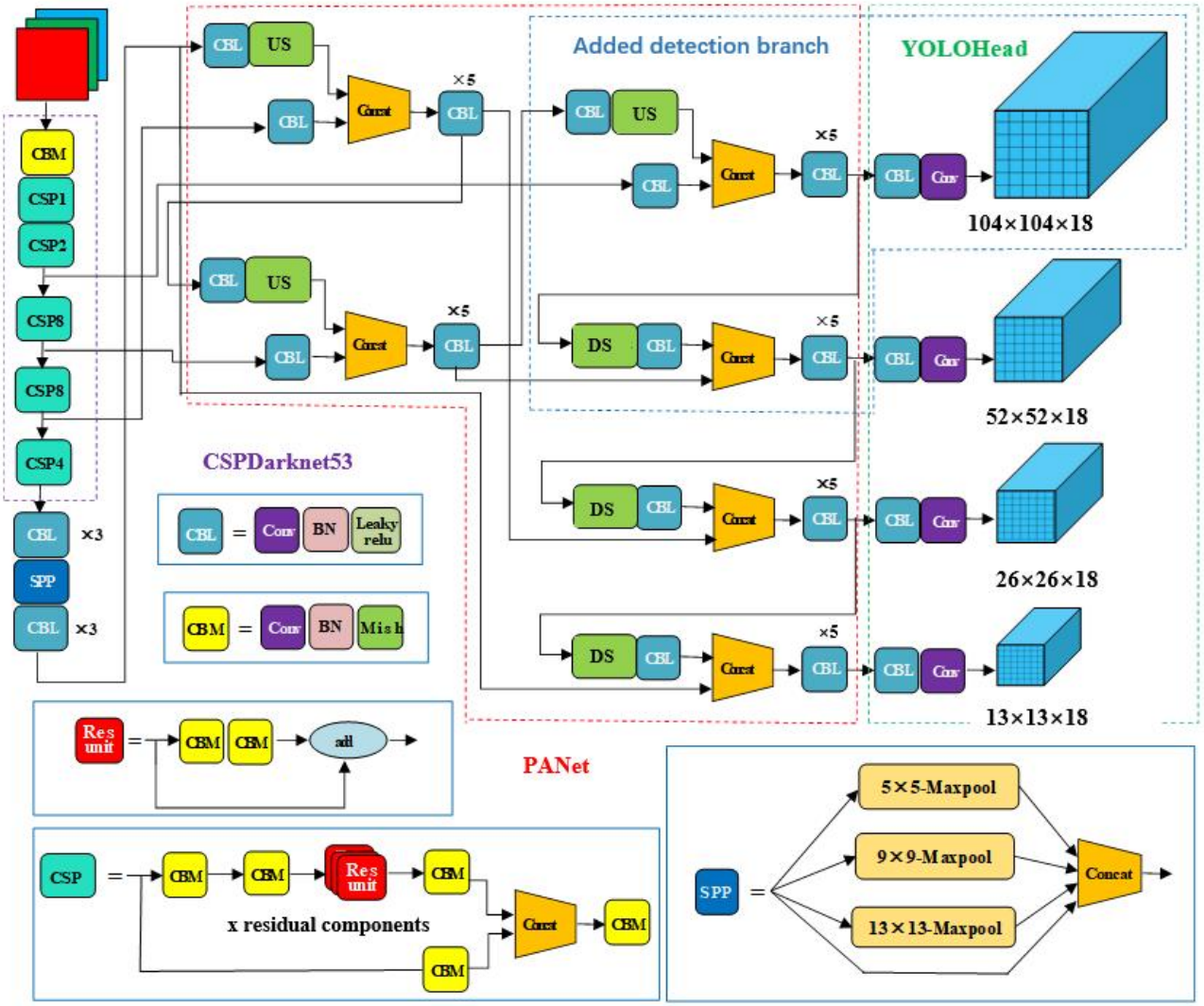


Fig. 2. *Block diagram of YOLOv4 structure.*

infrared images can further enhance the robustness of the model. Expansion of the dataset using data augmentation methods can lead to a more paradigmatic and accurate trained model .

In this paper, we do Physical space change and Color space transformation on the image.The methods and parameters used to perform data enhancement on the images in this paper are shown in Table I.We expanded the dataset through data augmentation to ensure a better fit of the model.

2.2. B. Improved YOLOv4

We propose a method for infrared image recognition of faulty areas of electrical equipment based on an improved YOLOv4 model, where the model structure is shown in Fig. 2. The basic idea of object detection in the YOLO model is to divide the image into a series of grids, traverse each grid, draw the frame if the detected grid contains an object, and output the final target prediction frame after traversal. Object classification and regression of prediction frames are performed simultaneously, and the whole process takes only one step. As shown in Fig. 2, the YOLOv4 model mainly consists of an input layer, a backbone layer, a neck layer and an output layer. The input layer receives fixed-size images, and after extracting features through the backbone network, it performs upsampling and feature aggregation in the neck layer, and then feeds them to the YOLO-Head classifier to output prediction frames at different scales, and finally, the class, confidence level and rectangular frame position of the target are obtained by frame regression.

The backbone feature extraction network of YOLOv4 uses csparknet53, which mainly consists of a CBM convolution module and a CSP residual module. the CBM convolution module consists of a convolution layer, a normalization layer and a Mish activation function layer. YOLOv4 chose the Mish function as the activation function of the CBM convolution module with the expression:

$$\text{Mish} = x \times tanh\left(ln\left(1 + e^{x}\right)\right) \quad (1)$$

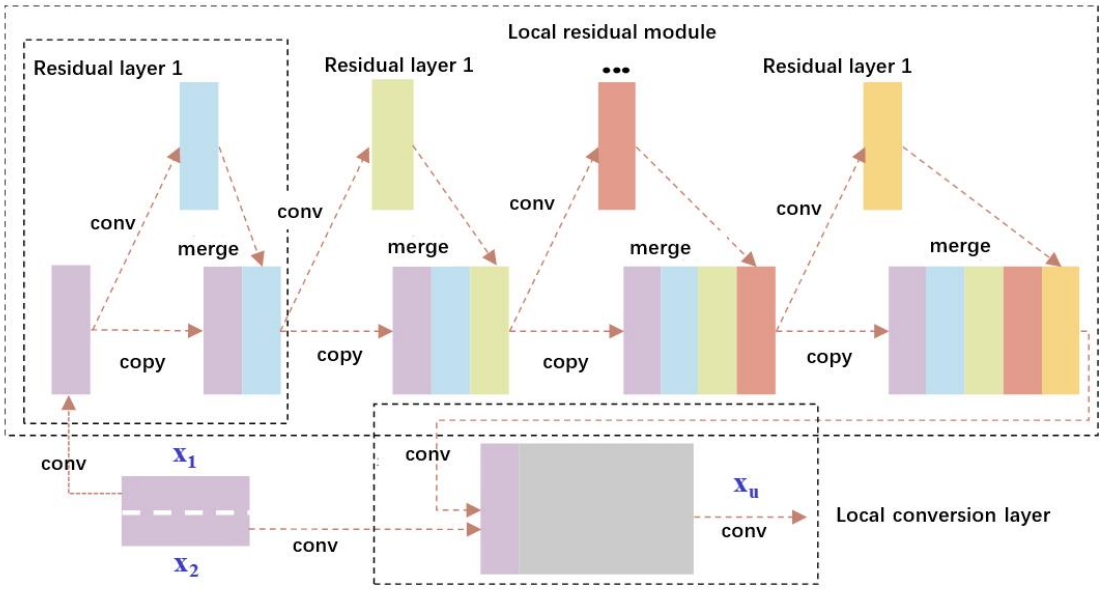


Fig. 3. *Cross-stage local residual networks.*

The Mish function has good smoothness, allowing small negative gradients to flow in and ensuring the integrity of the information flow.

The structure diagram of the cross-level local residual network is shown in Fig. 3. The structure divides the convolutional features input from the upper layer into two parts in a certain proportion, one part is used to superimpose the convolutional blocks and the other part is used for the convolutional calculation of large residual edges. Each time the newly combined features are convolved and the generated new features are merged, and finally the final features are output through the local transformation layer. The CSPnet mechanism can reduce the repetition rate of gradient information during the backpropagation of the network, further reduce the computation and improve the learning ability of the convolutional neural network.

The structure diagram of the cross-level local residual network is shown in Fig. 3. The structure divides the convolutional features input from the upper layer into two parts in a certain proportion, one part is used to superimpose the convolutional blocks and the other part is used for the convolutional calculation of large residual edges. Each time the Fig. 4 shows the Path Augmentation Network (PANet) architecture. The path aggregation network layer (PANet) combines a feature pyramid network and a path enhancement network. The feature pyramid network effectively propagates the high-level localization information from top to bottom, making localization more accurate. The path augmentation network propagates the underlying semantic information from the bottom up, which facilitates classification.

The convolution layers of the YOLOv4 model are as high as 60 or 70 convolutions. If we follow the up and down sampling method of the common feature pyramid, the bottom layer information flows to N2~N5, which requires several convolutions and the information flow line is too long. With the green dashed line in Fig. 4, the bottom features can be made to flow to N2~N5 through only a few convolution layers, which greatly improves the computational speed and feature utilization.

The YOLO head is a classifier consisting of a CBL convolution module, convolution and regression. The original YOLOv4 model extracts three feature layers of the backbone network for target detection, which are located in the middle layer, lower middle layer, and bottom layer. Their sizes and dimensions are (52, 52, 256), (26, 26, 512), and (13, 13, 1024), respectively. The YOLOv4 model predicts three prior boxes for the feature maps of each feature layer. Each prior box includes: the position of the box (coordinates of the center point, height and width of the box), the category of the target object, and N categories.

Among the thermal anomaly of electrical equipment, the most common is the current heating type fault, which occurs mainly at fixtures and connection wires. The thermal area is small, and the overheating area shown in the image accounts for a small part of the whole image. In order to accurately detect small hot spots of electrical equipment, this paper makes corresponding improvements based on the original YOLOv4 model. The improvement part is shown in Fig. 2. The improvement strategy is 104 × 104, and a new target detection branch is added to the shallow feature map of 104. The feature layer is obtained by four times downsampling, which better preserves the semantic information of small targets. The improved YOLOv4 model is changed from the original 13×13, 26×26, and 52×52 three-scale detection layers to 13×13, 26×26, 52×52, and 104×104.

Fig. 5 shows a schematic representation of the different sizes of the a priori frames on the feature map. It can be seen that the feature map is divided into large receptive fields with a small number of grids and the corresponding a priori frames are large in size, suitable for detecting large target objects. On the contrary, the feature map is divided into large receptive fields with a large number of grids, and the corresponding a

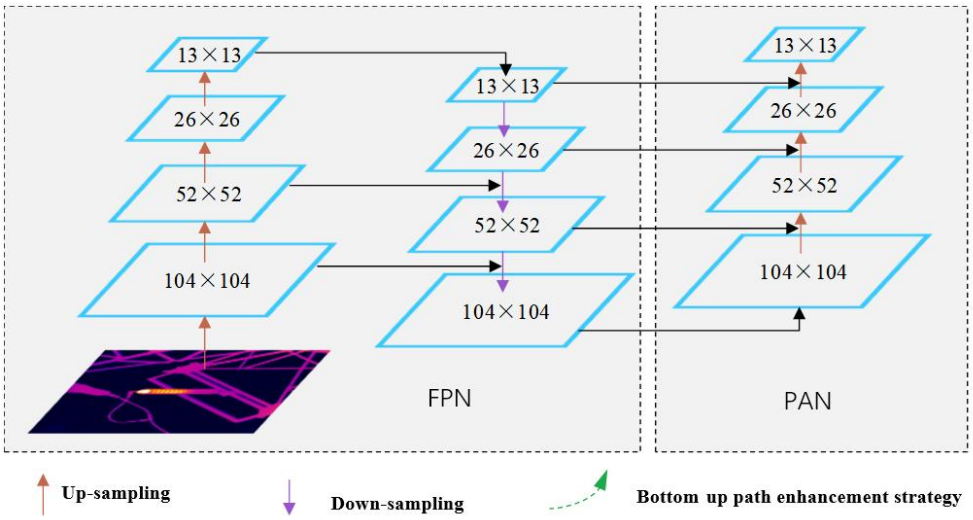


Fig. 4. *Path Augmentation Network (PANet) architecture.*

priori frame size is smaller, suitable for detecting small target objects. For heating faults of the image converter, the larger the a priori frame, the larger the area of the detected fault region, while the smaller the a priori frame, the better the fault region can be enclosed.

Compared with the original YOLOv4 model, this paper has an additional requirement of 104 × 104 to assign an appropriate prior frame size for the feature map. By observing and analyzing the detection objects in this paper, K-means clustering of real frames from the training set is performed in this paper to obtain an a priori frame size that is consistent with hot fault zone detection.

Compared with the original YOLOv4 model, this paper has an additional requirement of 104 × 104 to assign an appropriate prior frame size for the feature map. By observing and analyzing the detection objects in this paper, K-means clustering of real frames from the training set is performed in

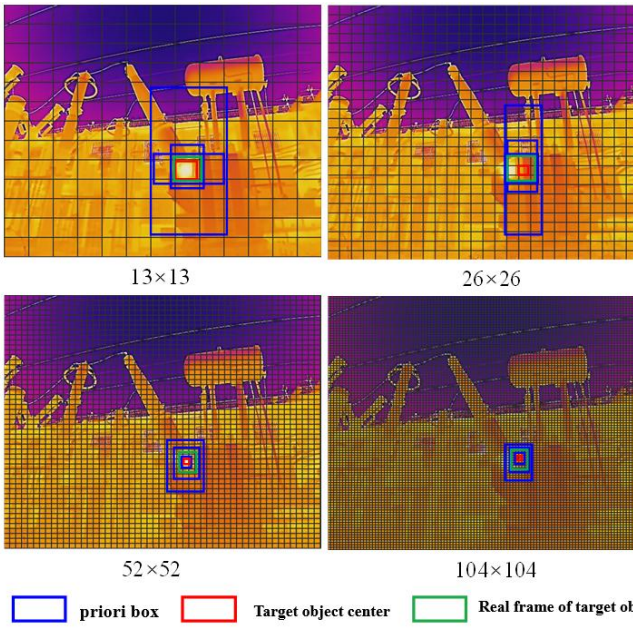


Fig. 5. *A priori frame on feature map.*

this paper to obtain an a priori frame size that is consistent with hot fault zone detection.

The basic idea of the K-means algorithm is to select K data points as the initial clustering centers for clustering dimensional space data, calculate the Euclidean distance between all data and the initial clustering centers, classify the objects closest to the centers, and continuously update the clustering centers by iterative methods until the sum of all Euclidean distances is minimized. The schematic diagram of this algorithm is shown in Fig. 6.

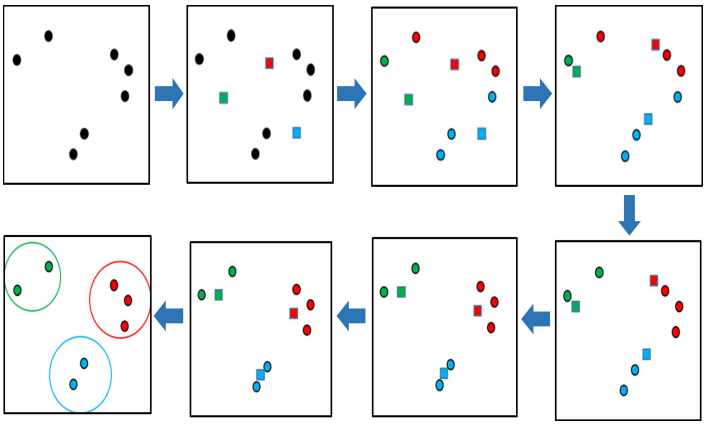


Fig. 6. *K-means clustering algorithm.*

For the improved four-scale YOLOv4 detection model in this paper, three prior frames are required for each detection scale, so the number of clusters in the K-means clustering algorithm is set to 12. After clustering the actual frames of the training set data, the sizes of the four sets of prior frames obtained are shown in Table II. The results show that the corresponding a priori frames of feature maps divided into fewer grids are also larger in size, and conversely, the corresponding a priori frames of feature maps divided into more grids are smaller in size.

TABLE II
PRIOR FRAME SIZE OBTAINED BY K-MEANS CLUSTERING

Feature map	13×13	26×26	52×52	104×104
priori box	(68, 83) (135, 57) (155, 304)	(44, 149) (46, 50) (67, 216)	(25, 26) (29, 41) (32, 85)	(9, 11) (13, 15) (17, 21)

In the training of the YOLOv4 model, this chapter employs a unique mosaic enhancement technique. During the training process, four images from a certain batch of samples are randomly selected and stitched into new images by random scaling, random cropping and random alignment. As shown in Fig. 7, one image of the mosaic-enhanced data will contain multiple hot-fault regions. The mosaic-enhanced image is more beneficial for detecting small target thermal fault regions than the previous case where the number of thermal fault regions was small and numerous.

All categorical labels in the YOLO model are encoded with either 0 or 1. During the training process, the position will output 1 if the model detects that the target belongs to the nth category, otherwise it will output 0. The coding of the classification labels is in the form of what is called one-time hot coding. Due to the limited data set, too much training time can easily lead to overfitting. This problem can be solved by using label smoothing techniques.

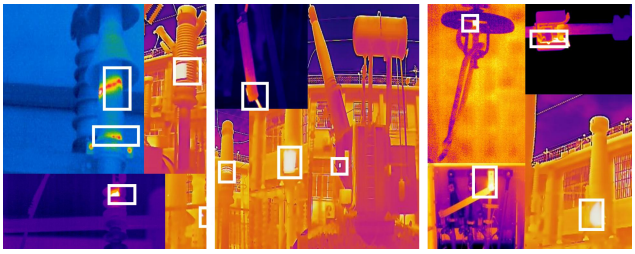


Fig. 7. *Partial mosaic enhanced data.*

Label smoothing is a regularization strategy that "softens" the labels obtained by traditional single-thermal coding by reducing the weights of the real samples when calculating the loss function, thus preventing overfitting when calculating the loss values:

$$\begin{cases} y_i' = (1-\varepsilon)\cdot y_i + \varepsilon u(K) \\ u(K) = \dfrac{1}{K-1} \end{cases} \tag{2}$$

3. Experiments and results analysis

3.1. Experiment of Training Strategies

The experimental platform was a windows 10 64-bit system, the programming language was Python 3.7, and the deep learning code environment was Pytorch. Experiments were conducted on a desktop computer with an Intel(R) core(TM) i5-10400f CPU 2.90ghz and an NVIDIA Ge force RTX 2060 GPU. The models are trained using GPU training mode. Before training, the original 719 IR images were randomly divided into training dataset, validation dataset and test dataset in the ratio of 8:1:1, and the size of all images were normalized to 416 × 416 In this paper, 100 epochs were trained for the constructed model.

The loss function is a measure of the consistency between the predicted and actual results of the YOLO model, and further adjusts the parameters of the network training according to the change of the loss value. When the YOLO model is used for target detection, the output includes the bounding box of the target location, the confidence of the bounding box and the target type, so the corresponding loss function also has three parts.

The loss function in YOLOv4 consists of three parts:

$$Loss = Loss_{location} + Loss_{confidence} + Loss_{class} \tag{3}$$

$Loss_{location}$ is the frame position loss, $Loss_{confidence}$ is the frame confidence loss, and $Loss_{class}$ is the target category loss. The punishment factor added by YOLOv4 takes into account the aspect ratio of the prediction frame and the target frame, and uses *CIoU*. The calculation formula of *CIoU* is:

$$CIoU = IoU - \frac{\rho^2(b, b^g)}{d} - \alpha v \tag{4}$$

α and v are calculated as:

$$\alpha = \frac{v}{(1-IoU)+v} \tag{5}$$

$$v = \frac{4}{\pi^2}(\arctan\frac{w^g}{h^g} - \arctan\frac{w}{h})^2 \tag{6}$$

IoU is the intersection and union ratio of the predicted bounding box and the real bounding box, b is the center point of the predicted bounding box, b^g is the center point of the real bounding box, $\rho^2(b, b^g)$ is the Euclidean distance between the predicted bounding box and the center point of the real bounding box, d is the diagonal distance of the minimum rectangular region that can contain both the predicted bounding box and the real bounding box, and w^g, h^g, w and h are the width and height of the real bounding box and the predicted bounding box respectively.

This chapter uses freeze training to train the model. Freeze training incorporates the idea of migration learning, i.e., pre-training the network and then saving it for further training of subsequent networks. In this chapter, during model training, the first 50 epochs are frozen, the learning rate is set to 1 x 10-3, and the number of samples in the batch is set to 14. The last 50 epochs are thawed for training, the learning rate is set to 1 x 10-4, and the number of samples in the batch is set to 2.

When training the model, the training loss value and the validation loss value are saved after each epoch element in order to fine-tune the parameters of the model by observing the change in the loss value. To verify the effectiveness of the cosine annealing algorithm in tuning the learning rate, the

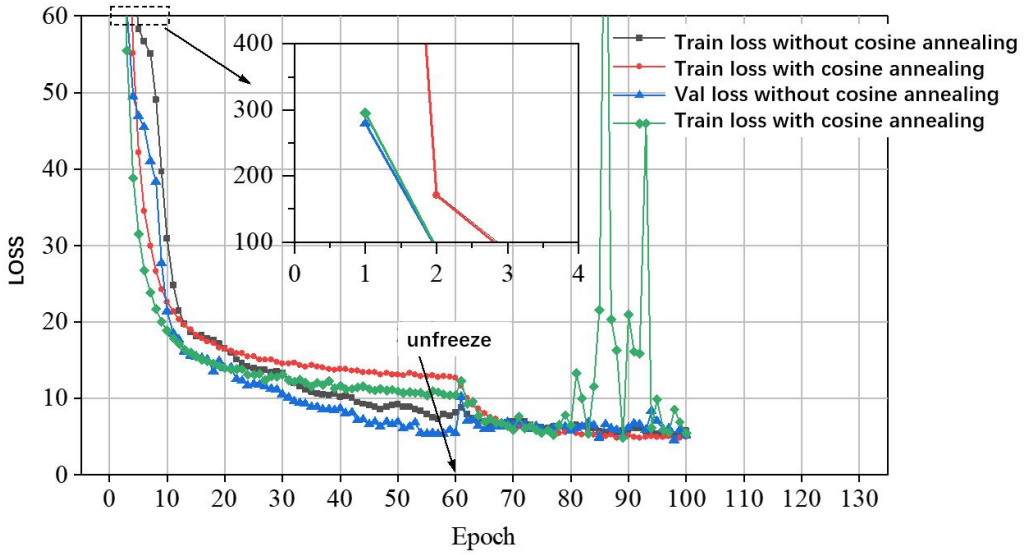


Fig.8. *Comparison of training and validation loss values with and without cosine annealing algorithm.*

TABLE III
THE MODEL PERFORMANCE BY USING DIFFERENT STRATEGIES TO IMPROVE THE NETWORK

Model Training Strategies	AP@0.5	DA	Pr	Re	F1	MR-2
Triple detection layers	30.94%	71.43%	0.61	0.22	0.32	0.82
Four detection layers	42.92%	70.94%	0.67	0.31	0.42	0.73
Four detection layers + K-means	51.92%	85.2%	0.62	0.43	0.53	0.68
Four detection layers + K-means + cosine annealing	59.53%	83.74%	0.80	0.26	0.39	0.61
Four detection layers + K-means + mosaic data enhancement	76.77%	92.61%	0.87	0.59	.0.70	0.42
Four detection layers + K-means + mosaic data enhancement + cosine annealing	83.17%	93.10%	0.78	0.872	0.7044	0.33

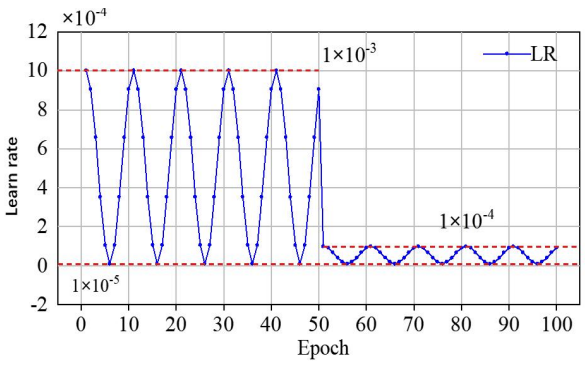


Fig.9. *Learning rate variation under cosine annealing algorithm.*

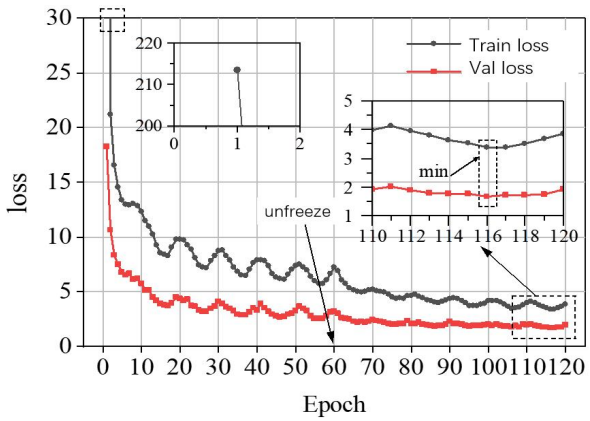


Fig.10. *Model training and validation of loss values.*

model training results are compared here, as shown in Fig. 8. It can be seen that before thawing, the training and validation loss values using cosine annealing are smaller than the values under normal conditions and there is no decay. After thawing, the validation loss values of the model trained without cosine annealing fluctuate and the model appears to be overfitted, indicating that the learning rate is too high at that stage. It can be seen that using cosine annealing to control the learning rate is beneficial for model convergence. Fig. 9 shows the variation pattern of the learning rate under the cosine annealing algorithm.

To verify the effectiveness of the improved YOLOv4 method in this paper, we used different strategies for model training and testing, and the obtained metrics are shown in Table III.

3.2. Anomaly area identification results and Comparison

Based on the original data images, one of the image enhancement methods listed in Table I was added as the total data set for model training, validation and testing, and each metric obtained by adding each image enhancement method to the test set is shown in Table IV.

It can be seen that the training loss value decreases rapidly from 213 for the first training to 21 for the second training, and the overall loss value and validation loss of the model gradually decrease and stabilize as the number of training increases, indicating that the model is converging. Before freezing the training, the training is unstable due to the larger learning rate and fewer model parameters at this time, making the training and validation loss values slightly more volatile. After thawing, the learning rate is smaller, the model parameters are larger, and the loss values change less and gradually stabilize. The model with the minimum training loss and validation loss values obtained in the 116th training was used as the final fault identification model.

The 145 infrared images of electrical equipment containing thermal faults are put into the model for testing, and then the model outputs the location and target confidence level of the rectangular box of the faulty region. The infrared image recognition results of the fault region of electrical equipment under different cases are shown in Fig. 11.

Fig. 11 shows the results of infrared image recognition of fault areas of electrical equipment in outdoor substations under different situations. Fig. 11(a) and Fig. 11 (c) show the detection of individual equipment in the substation where

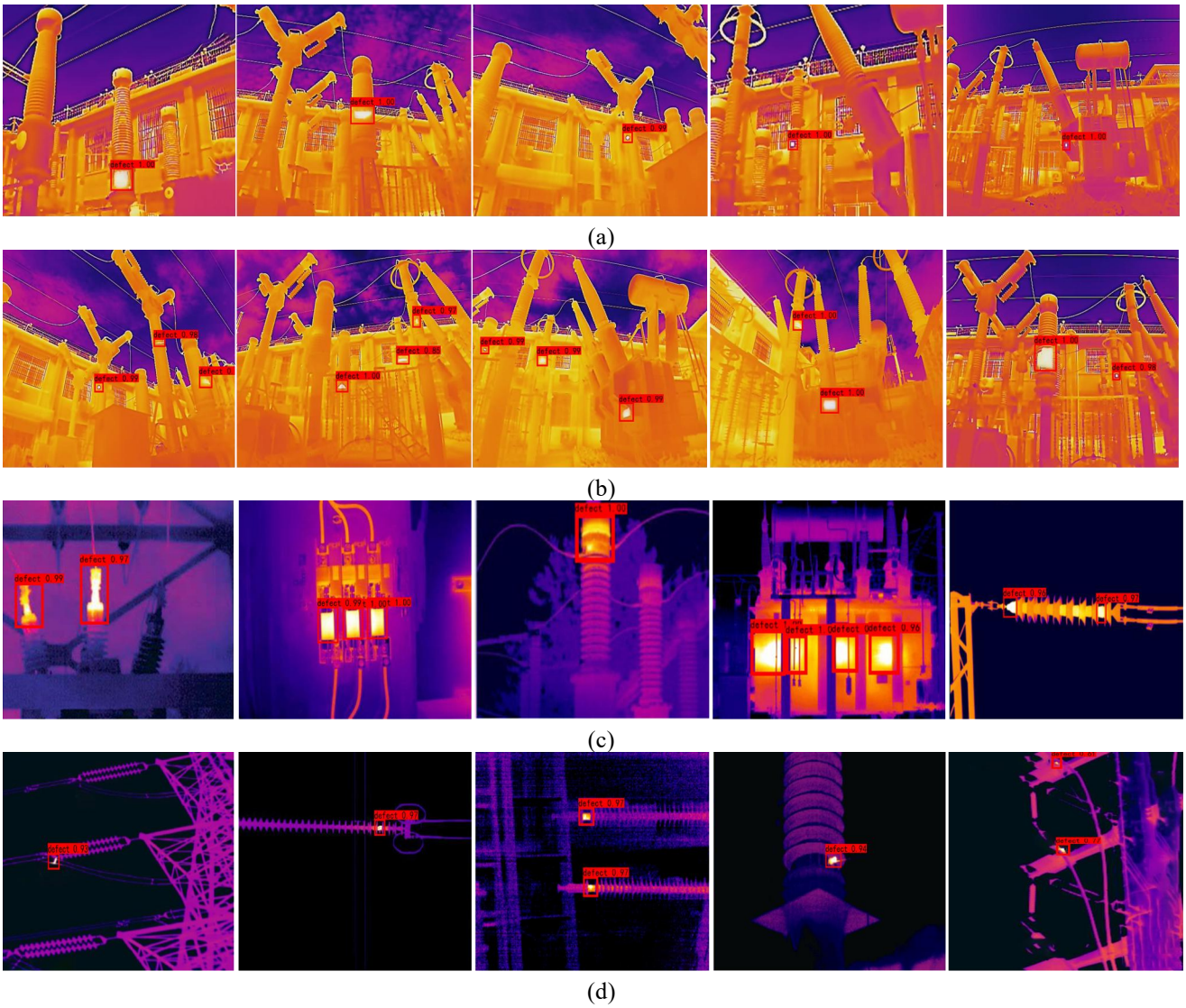


Fig. 11. *Results of infrared image recognition of faulty areas of electrical equipment. (a) Individual faults exist in different devices. (b) Multiple faults exist in the same scenario. (c) Different color palettes. (d) Small target hot zone.*

TABLE IV
RESULTS OF FAULT AREA IDENTIFICATION

Number to be identified	Number of models identified	MAP	Detection speed(f/s)	Pr	Re	F1	AP@0.5	MR-2
989	958	96.78%	17	0.92	0.92	0.93	94.23%	0.11

the heating area exists. It can be seen that the model can well detect the fault areas existing in current transformers, voltage transformers, circuit breakers, surge arresters and transformers, and the confidence of fault area detection are close to 1. And for small area fault areas this model can also accurately detect them. Fig. 11 (b) shows the detection results of a graph containing multiple devices with heat-emitting areas. From the results, it can be seen that the model in this paper is able to identify accurately regardless of the size and number of heat-emitting regions. Fig. 11(d) shows the detection results for a small heat dissipation area. It can be seen that the model has better detection for small targets, which is important for thermal fault detection during infrared imaging of line insulators and tower fixtures at long distances.

The calculated results of the indicators obtained from the model in this paper are shown in Table IV. The accuracy of the model is as high as 96.78% and the detection speed is 17 f/s. AP@0.5 is 94.23%, indicating that the accuracy and detection speed of the model are high and the model has good performance in electrical equipment fault detection.

To further validate the superiority of the model proposed in this paper for infrared identification of electrical equipment faults, different feature extraction networks corresponding to other current fault identification models were used to compare the test results and calculate each metric, and the results are shown in Table V.

From Table V, it can be seen that the AP@0.5 obtained by Faster R-CNN for fault region identification measurements is lower than the other algorithms, and the MR-2 value is also high, indicating that this series of algorithms is less effective in detecting small targets. In terms of detection speed, the SSD algorithm has the fastest detection speed, the YOLO series algorithm is the second, and the Faster R-CNN has the slowest detection speed. The highest AP@0.5 of 94.23% is obtained by this algorithm, which is 6.52% better than the original YOLOv4 model, and 4.67% better in terms of accuracy. the highest accuracy of 97.88% is obtained by the SSD (VGG16) model, which is 1.1% higher than the model in this paper, but the AP@0.5 of this model is 3.82% higher than it, and the MR-2 is also lower than it 0.09, so the recognition performance of the model in this paper is better. In summary, the model in this paper has certain superiority in performance for fault infrared recognition.

Aiming at the problems of irregular heating regions of electrical equipment, small heating targets and low detection accuracy, an improved YOLOv4-based infrared image recognition algorithm for fault regions of electrical equipment is proposed, and the fault region recognition algorithm is studied by constructing a data set for model training and model testing. Model testing results show that the improved four-detection layer model improves AP by about 12% compared with the original three-detection layer model. Model robustness test experimental results show that the model robustness is good.

3.3. Model Robustness Testing

In order to verify the robustness of the improved YOLOv4 fault region recognition model, robustness test experiments with Gaussian noise, pretzel noise, luminance variation, and blur variation are performed on 719 test set images.

For Gaussian noise, the mean value is set to 0, the variance is set to 0.05~0.15, and the interval size is 0.01. For the salt and pepper noise, the density percentage is changed from 0.01 to 0.1, the interval size is 0.01, and the image brightness is changed to 50% to 150% of the original brightness, with an interval of 10%. Each of these three methods is applied to the test dataset to obtain a new dataset, and the transformed dataset is fed into the model for testing.The test results are shown in Figure 12.

Fig.s 12(a) and 12(b) show that the AP and DA metrics of the model still exceed 80% under low noise, indicating that the model is robust to noise disturbances. However, the model performance degrades to a certain extent when the noise is high, and it is necessary to ensure that the images have good

TABLE V
COMPARISON OF FAULT AREA IDENTIFICATION RESULTS OF DIFFERENT MODELS

Models	AP@0.5	DA	F1	Precision	Recall	MR-2	FPS
Faster R-CNN(VGG16)	77.79%	82.08%	0.58	0.429	0.912	0.41	10
Faster R-CNN(Restnet50)	81.12%	83.08%	0.59	0.432	0.923	0.35	5
SSD(VGG16)	90.41%	92.98%	0.58	**0.961**	0.413	0.20	**28**
SSD(Mobiletnet)	85.95%	86.35%	0.78	0.690	0.897	0.28	23
YOLOV3(Darknet)	89.57%	93.12%	0.92	0.919	0.920	0.17	16
YOLOv4(CSPDarknet)	87.71%	88.95%	0.87	0.908	0.837	0.22	13
Our Model	**94.23%**	**96.78%**	**0.93**	0.923	**0.926**	**0.11**	17

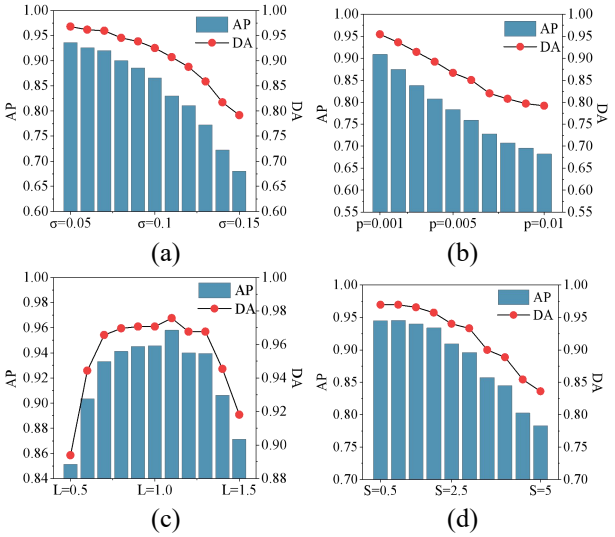


Fig. 12. *Robustness testing results. (a) Gaussian noise. (b) Salt And Pepper Noise. (c) Brightness. (d) Gaussian Blur.*

quality. From Fig. 4.12 (c) it is shown that a decrease or increase in luminance relative to the original image causes the device AP@0.5 and DA to decrease, both values varying within a small range and above 80% overall, indicating good robustness of the model to changes in luminance. Fig. 4.20 (d) shows that as the standard deviation of the Gaussian kernel function increases, the image becomes increasingly blurred and the AP@0.5 and DA of fault identification both decrease, but the model accuracy remains high in a small range, indicating that the model is robust to Gaussian blur.

4. Conclusion

In this paper, We propose an improved YOLOv4 algorithm for infrared image anomaly area identification, which can detect the thermal generation phenomenon of electrical equipment. a new target detection branch based on the shallow 104×104 feature map is added to the model of the original YOLOv4, which solves the problems of irregular thermal area of electrical equipment, small thermal targets and low detection accuracy. The improved four-detection layer model improves about 12% AP@0.5 over the original three-detection layer model .

We constructed a dataset containing 719 images. We used infrared cameras to capture images of various abnormal equipment in the substation and collected some images from the Internet to expand the dataset. To make the training better we used data augmentation on the dataset. During model training, three training techniques, mosaic data enhancement, label smoothing, and cosine annealing, are used to improve the accuracy of the training results. Compared with the original model, the AP value of the resulting model improves by 52.23% and the accuracy improves by 21.67% after improving the model using the three methods.The comparison with other models shows that this model has excellent detection effect on infrared images of electrical equipment faults.

The robustness of the improved model is tested, and the results show that the model has good robustness to noise, luminance and blur. The model is resistant to a certain level of interference and can be adapted to different situations in the work.

5. REFERENCES

[1] Zheng, Hanbo, et al. "Infrared image detection of substation insulators using an improved fusion single shot multibox detector." *IEEE Transactions on Power Delivery* 36.6 (2020): 3351-3359.

[2] Osornio-Rios, Roque Alfredo, Jose Alfonso Antonino-Daviu, and Rene de Jesus Romero-Troncoso. "Recent industrial applications of infrared thermography: A review." *IEEE transactions on industrial informatics* 15.2 (2018): 615-625.

[3] Ullah, Irfan, et al. "Predictive maintenance of power substation equipment by infrared thermography using a machine-learning approach." *Energies* 10.12 (2017): 1987.

[4] Huang, Song, et al. "Research on Real-Time Disconnector State Evaluation Method Based on Multi-Source Images." *IEEE Transactions on Instrumentation and Measurement* 71 (2021): 1-15.

[5] Han, Sheng, et al. "A smart thermography camera and application in the diagnosis of electrical equipment." *IEEE Transactions on Instrumentation and Measurement* 70 (2021): 1-8.

[6] Usamentiaga, Ruben, et al. "Temperature monitoring for electrical substations using infrared thermography: architecture for Industrial Internet of Things." *IEEE transactions on industrial informatics* 14.12 (2018): 5667-5677.

[7] Huda, A. S. N., et al. "A new thermographic NDT for condition monitoring of electrical components using ANN with confidence level analysis." *ISA transactions* 53.3 (2014): 717-724.

[8] Li, Yongbo, et al. "A new intelligent fault diagnosis method of rotating machinery under varying-speed conditions using infrared thermography." *Complexity* 2019 (2019).

[9] Jadin, Mohd Shawal, Soib Taib, and Kamarul Hawari Ghazali. "Feature extraction and classification for detecting the thermal faults in electrical installations." *Measurement* 57 (2014): 15-24.

[10]Luo, Yanhong, et al. "A survey of intelligent transmission line inspection based on unmanned aerial vehicle." *Artificial Intelligence Review* (2022): 1-29

[11]Hao, Yanpeng, et al. "Methods of image recognition of overhead power line insulators and ice types based on deep weakly-supervised and transfer learning." IET generation, transmission & distribution 16.11 (2022): 2140-2153.

[12]Mishra, Manohar, et al. "Deep learning in electrical utility industry: A comprehensive review of a decade of research." *Engineering Applications of Artificial Intelligence* 96 (2020): 104000.

[13]Li, Bing, et al. "Two-Level Model for Detecting Substation Defects from Infrared Images." Sensors 22.18 (2022): 6861..

[14]Shao, Haidong, et al. "Intelligent fault diagnosis of rotor-bearing system under varying working conditions with modified transfer convolutional neural network and

thermal images." *IEEE Transactions on Industrial Informatics* 17.5 (2020): 3488-3496..

[15] Zhang, Qianyi, et al. "Equipment detection and recognition in electric power room based on faster R-CNN." *Procedia Computer Science* 183 (2021): 324-330.

[16] She, Lingcong, et al. "Insulator surface breakage recognition based on multiscale residual neural network." *IEEE Transactions on Instrumentation and Measurement* 70 (2021): 1-9.

[17] Ou, Jianhua, et al. "Infrared Image Target Detection of Substation electrical equipment Using an Improved Faster R-CNN." *IEEE Transactions on Power Delivery* (2022).

[18] Wang, Shenghui, Leilei Niu, and Nan Li. "Research on image recognition of insulators based on YOLO algorithm." *2018 international conference on power system technology (POWERCON)*. IEEE, 2018.

[19] Lianqiao, Li, et al. "Recognition and application of infrared thermal image among power facilities based on YOLO." 2019 Chinese Control And Decision Conference (CCDC). IEEE, 2019.