International Journal for Housing Science and Its Applications
Publish August 5, 2025. Volume 46, Issue 3 Pages 4387-4398

https://doi.org/10.70517/ijhsa463375

Research on safety monitoring model of substation operation environment based on remote sensing technology

Hengjie Liu[1,*], Jie Tang[2] and We Li[1]
[1] Zhengzhou Electric Power Co., Ltd., Zhengzhou, Henan, 450064, China
[2] Department of Mechanical Engineering, Henan University of Science and Technology, Zhengzhou, Henan, 450064, China
Corresponding authors: (e-mail: L15333193316@126.com).

Abstract With the expansion and complexity of the power system, the safety monitoring of the substation operation environment has become a key link to ensure the life safety of power operators and the stable operation of equipment. This study proposes a safety monitoring method for substation operation that integrates high resolution remote sensing technology and improved YOLOv5 deep learning model, and constructs a complete monitoring system through three core modules: remote sensing image preprocessing, target detection algorithm optimization about safety equipment and dynamic detection of dangerous areas. In remote sensing image segmentation, a multi-threshold segmentation method is used to eliminate geometric distortion and radiation distortion and extract key feature information. For the problem of small target detection in complex scenes, the YOLOv5 model is improved, the coordinate attention mechanism CA is embedded to enhance the feature extraction capability, and the SPPF module is reconstructed by using the large kernel separated attention convolution, which is combined with the GD aggregation-distribution mechanism to optimize the necking network and to improve the multi-scale target detection accuracy. The experimental results show that the improved model has a denoising performance PSNR of 28.98dB, an SSIM of 0.874, and a safety equipment detection F1 value of 95.10% for insulated suits and 95.55% for safety helmets. The average accuracy of hazardous area misentry detection is 95.85%, etc. are significantly better than YOLOv3, Faster RCNN and other comparative models, and the computational efficiency is high, and the detection speed reaches 7.26ms/mg.

Index Terms high resolution remote sensing, substation operation, safety monitoring, YOLOv5, hazardous area misentry detection

I. Introduction

Electricity is the most indispensable kind of energy in human production and life, and it plays a certain role in promoting the process of social development. Substation is a place used for changing voltage, exchanging power and distributing electric energy in the electric power system, which is mainly composed of transmission and distribution, transformer, auxiliary control, operation and maintenance, etc. [1]. In order to provide a guarantee for the smooth development of production and life work, and also to provide a guarantee for the smooth implementation and stability of power transmission, the operation and maintenance of the substation is a key link to ensure the safe and stable operation of the power grid [2], [3]. At present, the field operation safety management requirements have been meticulous, but by the environmental conditions, personnel quality, protection technology and other factors, substation maintenance field operation in the implementation of more safety risks still exist [4], [5]. Reasonable protective measures need to be taken to reduce the probability of personnel liability accidents, and promote the sustainable and healthy development of substation operation and maintenance work [6].

In the specific implementation of substation operation and maintenance work process, the adoption of effective safety protection and monitoring technology, will help the work better implementation, in reducing equipment maintenance cost investment at the same time to ensure that the user can get a stable power supply [7]-[9]. The application of modern monitoring facilities in the process of carrying out specific substation operations, through the adoption of intensive monitoring mode to analyze the operating environment, can effectively improve the quality of substation operation and maintenance work, improve the professionalism and efficiency of substation operations [10]-[12]. Through the implementation of strict safety protection measures and monitoring mechanisms, the safety condition of the operation site can be monitored in real time, and potential safety hazards can be discovered and corrected in a timely manner, thus effectively preventing accidents and safeguarding the lives of employees [13], [14]. At the same time, the combination of monitoring and collecting data and fault formation conditions, relying on the role of the intelligent system should be developed to improve the corresponding countermeasures, can provide regulatory reference for the management department to better carry out their work [15], [16]. Therefore, regarding the substation operation and maintenance field operation should strengthen the safety protection, improve the implementation of monitoring technology attention, in order to protect the normal functioning of substation equipment [17].

This study proposes a safety monitoring method for substation operation that integrates high-resolution remote sensing technology and deep learning model, aiming at realizing efficient and accurate operation environment analysis and dynamic detection of dangerous areas. The article is centered on three core areas: remote sensing image processing, safety equipment detection and hazardous area misentry detection. High-resolution remote sensing image segmentation eliminates geometric distortion and radiometric distortion through data preprocessing, and extracts key feature information by combining with threshold segmentation method to provide high-quality data basis for subsequent analysis. Deep learning-based safety equipment detection model for substation operation: for the complex scene of substation, the YOLOv5 target detection algorithm is improved, the sample diversity is expanded by online data enhancement, and the coordinate attention mechanism CA is embedded in the backbone network, which enhances the positional sensitivity and feature expression of the insulated gloves and helmets by decomposing the channel and encoding the positional information. And the SPPF module is reconstructed using large kernel separated attention convolution, combined with the GD aggregation-distribution mechanism to optimize the neck network, to solve the problem of shallow feature information loss, and to enhance the detection robustness of multi-scale targets. MPDIoU is also introduced as the bounding box regression loss function to reduce the matching error between the predicted and real boxes. Finally, a dynamic hazardous area misentry detection method is proposed, which maps the 3D safe area to the 2D imaging plane based on the single-stress transformation, determines whether the operator has misentered the hazardous area by combining the geometric rules, and realizes the real-time alarm function.

II. Safety monitoring of substation operations based on high-resolution remote sensing and deep learning

Users are the direct participants in realizing the remote sensing information-oriented transformer operation and play a crucial role in the system. Since the monitoring range of surface information is often wide and the remote sensing image of the corresponding area contains a huge amount of information, the multi-user division of labor is the most important operation mode. For substation operations with remote sensing information, the operation monitoring scheme determines the efficiency during substation operations. In this chapter, the substation operation monitoring scheme is investigated.

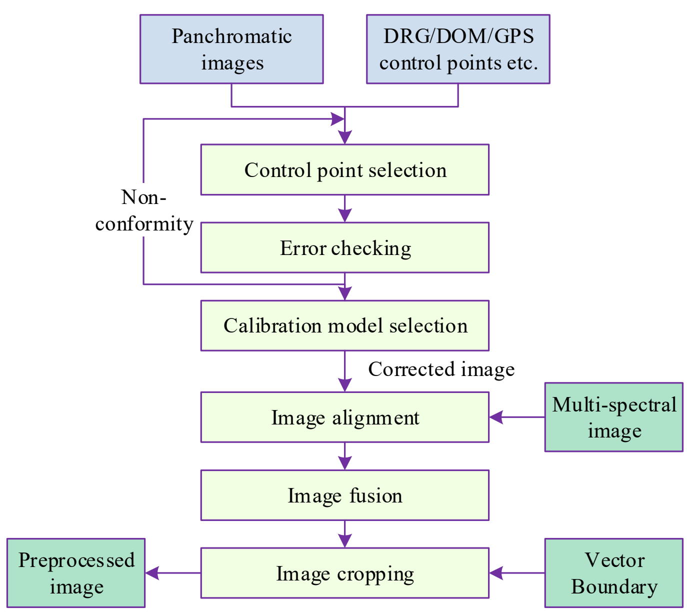

Figure 1: The general process of data preprocessing

II. A. Segmentation of high-resolution remote sensing substation operation images

II. A. 1) Pre-processing of remote sensing data

In the process of acquiring data, satellites are affected by satellite sensors and many other uncontrollable factors (terrain, climate conditions, shooting time, sun direction, etc.), which make the acquired remote sensing images have certain geometric distortions, radiation distortion, atmospheric extinction and other phenomena, and these distortions and aberrations affect the real value of the feature coverage. Remote sensing images need to be pre-processed accordingly to minimize these disturbing factors before performing change detection, segmentation, classification and other operations. Common data preprocessing methods include image alignment, radiometric correction, geometric correction, and so on. The general flow of remote sensing image data preprocessing about substation operations is shown in Figure 1.

II. A. 2) Threshold-based segmentation methods

Remote sensing data preprocessing effectively eliminates image distortion and noise interference, in order to further extract the key feature information, this section introduces the threshold-based segmentation method, which realizes the fine segmentation of the substation operation area through the dynamic selection of global and adaptive thresholds.

Threshold segmentation is the process in which the input image I is transformed by Equation (1) to obtain the output image O, and this type of method is most used in the field of image segmentation due to its advantages of simplicity, high operational efficiency, and high speed.

$$O(x,y)=\begin{cases}1, & I(x,y)\geq T\\ 0, & I(x,y)<T\end{cases} \tag{1}$$

where, T is the segmentation threshold; when the gray value of the pixel point in the input image I is greater than the threshold T, $O(x,y)=1$ and the pixel point is categorized as the foreground image element; when the gray value of the pixel point in the input image I is less than the threshold T, $O(x,y)=0$ and the pixel point is categorized as the background image element; x, y are the horizontal and vertical coordinates of the pixel point in the image.

The selection of the threshold value is the key to the threshold segmentation method. For different image characteristics and attributes, the main ones currently used are global thresholding, adaptive thresholding and optimal thresholding, etc.

Global thresholding refers to the use of the same threshold T for segmentation processing of the whole image, global thresholding segmentation methods only consider the size of the gray value of the pixel point to be processed, without considering the spatial information between the pixels in the field, suitable for images with obvious contrast between the gray value of the background and the foreground. When different parts of the same object on the same image have different brightness, the segmentation results are not satisfactory. Commonly used selection methods for global thresholding include peaks and valleys method, maximum inter-class variance method (OSTU), maximum entropy automatic thresholding method and so on.

Adaptive thresholding refers to the segmentation of an image by selecting different thresholds T1, T2...Tn in different regions according to the local features of the image, and it is used more in images where the contrast levels of the foreground and background gray values in the image are not uniform.

Optimal thresholding means that the thresholds are determined through experiments and specific problems for segmentation of high scores remote sensing images of substation operations.

II. B. Deep learning based safety equipment detection model for substation operation

Accurate segmentation of high-resolution remote sensing images provides basic data support for substation operation safety monitoring, however, the dynamic detection of small targets in complex scenes still needs to rely on more efficient algorithmic models. For this reason, this chapter further proposes a deep learning-based safety monitoring model for substation operation, which realizes the accurate identification of key targets by optimizing the network structure and attention mechanism.

II. B. 1) Algorithms for safety monitoring of transformer operations

Aiming at the difficulty of substation operation safety equipment wear detection, this paper carries out continuous improvement and optimization to enhance the detection performance of YOLOv5 target detection model, which has both high detection accuracy and fast inference speed. Firstly, online data enhancement is added to the data input of the model to enrich the number of samples in the data set and the sample background. Secondly, on the basis of the backbone feature extraction network, the CA module of coordinate attention mechanism is added to focus on the region of interest and enhance the feature extraction capability of the backbone network, and the SPPF module of spatial pyramid pooling is improved by using the large kernel separated attention convolution to enhance the multi-scale feature extraction capability. Again using GD aggregation-distribution mechanism to reconstruct and efficiently improve the neck structure through the deep fusion of deep and shallow feature maps, which effectively utilizes the shallow feature information and solves the loss of shallow feature information in the feature pyramid to enhance the feature fusion capability of the neck network. Finally, the use of MPDIoU as the model bounding box regression loss function efficiently improves the bounding box regression accuracy and reduces the model prediction error, thus enhancing the model detection accuracy.

II. B. 2) Feature Extraction Backbone Network Improvement

Due to the complex environment of the substation operation site and the large proportion of power operators, the introduction of the attention mechanism in the network can better obtain the important feature information of the insulated gloves and safety helmet small target objects, which can improve the small target detection accuracy. Compressed excitation (SE) and convolutional attention module (CBAM) are common attention mechanisms. Due to the small pixel points occupied by small targets, the target coordinate information is often biased during detection, but the SE and CBAM attention mechanisms only focus on the target feature information in the channel, using 2D global pooling, which results in the loss of the target position information and ignores the problem of the loss of position information for small targets. However, for capturing target structures, target position information is important.

Coordinate Attention Mechanism CA then decomposes channel attention into two parallel 1D feature encoding processes, one direction to obtain long-range dependencies, the other direction to obtain location information, aggregating the two directions feature information in encoding it, the encoded network model is sensitive to target localization and identification. The coordinate attention mechanism CA notices the spatially effective position and channel of the upper target feature information, suppresses the useless feature position information, enhances the small target detection accuracy, and improves the small target detection misdetection and omission problem. Therefore, this paper introduces the coordinate attention mechanism CA module on the basis of the improved YOLOv5 feature extraction backbone network. The structure of the coordinate attention mechanism CA is shown in Figure 2: The coordinate attention mechanism (CA) is a lightweight attention mechanism, different from the channel attention mechanism, the coordinate attention mechanism CA considers both the channel information and the position information, and embeds the target position information into the channel attention, so that the network model pays attention to more positional feature information in the region of interest.

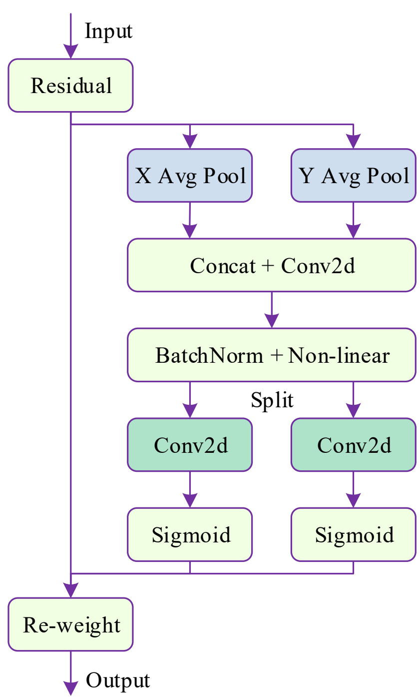

Figure 2: Coordinate A ttention Mechanism Architecture

The coordinate attention mechanism CA module can be used as a plug-and-play computational unit with enhanced feature extraction capability. The coordinate attention mechanism CA can take the output tensor $X=\left[X_1,X_2,\ldots,X_C\right]\in R^{C\times H\times W}$ is the input tensor, and an output tensor $Y=\left[Y_1,Y_2,\ldots,Y_C\right]$ with enhanced feature extraction capability is outputted by the Attention Mechanism module. Where C is the number of channels, H, W are the height and width of the input image.

The coordinate attention mechanism module contains two parts, one is coordinate information embedding and the other is coordinate attention generation. In the part of coordinate information embedding, in order to obtain the long-range dependence of the target's accurate location information, instead of using global average pooling, the global average pooling is decomposed into two one-dimensional feature encoding operations in two different directions, one using the pooling kernel size of $(H,1)$ in the direction of X, and the other one in the direction of Y, using the pooling kernel size of $(1,W)$. Operation, then for the input tensor X, the output height is h and the number of channels is c as shown in equation (2):

$$z_c^h(h)=\frac{1}{W}\sum_{0\le i\le W}x_c(h,i) \tag{2}$$

The output width is w and the number of channels is c as shown in Equation (3):

$$z_c^w(w)=\frac{1}{H}\sum_{0\le j\le H}x_c(j,w) \tag{3}$$

Input feature maps of size $C\times H\times W$ are pooled along X, i.e., horizontal, and Y, i.e., vertical, operations to obtain feature maps of size $C\times H\times 1$ and $C\times 1\times W$.

In the coordinate attention generation part, the resultant feature maps z^h, z^w output from the previous step are subjected to Concat operation, and the dimensionality reduction operation and activation operation i.e., F_t and δ are carried out on them by using the 1×1 convolution, and the output feature maps $f\in R^{C/r\times(H+W)}$ are outputted as shown in Eq. (4) shown:

$$f=\delta\left(F_t\left(\left[z^h,z^w\right]\right)\right) \tag{4}$$

where f is the feature map generated on the spatial information in X, i.e. horizontal direction, and Y, i.e. vertical direction, and r denotes the downsampling ratio. Then, f is split along the spatial dimension to generate X, i.e., horizontal direction, and Y, i.e., vertical direction, $f^h\in R^{C/r\times H\times 1}$ and $f^w\in R^{C/r\times 1\times W}$ feature maps, and the two feature maps are separately upgraded by 1×1 convolution, i.e., F_h and F_w, The number of channels is equal to the number of input tensor X channels, and the attention vectors $g^h\in R^{C\times H\times 1}$ and $g^w\in R^{C\times 1\times W}$ are obtained as attention weights by combining the sigmoid activation function, and the attention vectors are obtained as shown in equations (5) and (6):

$$g^h=\sigma\left(F_h\left(f^h\right)\right) \tag{5}$$

$$g^w=\sigma\left(F_w\left(f^w\right)\right) \tag{6}$$

Finally, the output of the CA module of the coordinate attention mechanism is shown in Equation (7):

$$y_c(i,j)=x_c(i,j)\times g_c^h(i)\times g_c^w(j) \tag{7}$$

The CA module is introduced as shown in Figure 3, the research in this paper mainly puts the coordinate attention mechanism CA module in the layer before the SPPF module, the ninth layer of the backbone network, at this time, the output feature map after the first eight layers of the backbone network not only has a clear contour and contains rich semantic feature information, the feature map after the output feature map of the CA module of the coordinate attention mechanism not only carries the target location information, but also contains the feature enhancement target feature information. This feature map is not only with target location information, but also contains feature information after enhancement, which is conducive to improving the feature extraction capability of the backbone network. The experiments in this paper have strongly proved that the CA module of the coordinate attention mechanism is effective in improving the detection accuracy of the substation operation safety monitoring.

Figure 3: Positional Introduction of CA Module

II. C.Hazardous area misentry detection methods

Although the target detection model based on deep learning can efficiently recognize operators and equipment, it is still necessary to determine whether they enter the hazardous area in real time in practical applications. In this chapter, a dynamic detection method of dangerous area misentry is proposed by combining the single responsive transformation and geometric rules to form a complete monitoring and control closed loop from target identification to safety determination.

In the process of electric power banding operation, the operator needs to strictly comply with the standardized operation requirements, and must maintain a certain safe working distance. In addition to relying on the operators to consciously and strictly comply with the operating norms, the virtual area can be delineated by the method of detecting the operator's accidental entry into the hazardous area, so as to timely detect potential safety hazards. According to the standard operating requirements to provide three-dimensional ground position information of the hazardous operation area (given a few coordinate points to uniquely determine a polygonal area), the single should be transformed and its linearity-preserving can be used to realize the delineation of hazardous areas in the two-dimensional pixel plane.

Firstly, it is necessary to determine the homology relationship between the actual ground area and the camera imaging plane, in order to ensure the accuracy of the monoresponse matrix, the method chooses a manual method to establish the world coordinate system in the ground area to be detected in advance, and selects several points evenly and labels them, so that several groups of marked points under different viewing angles can be obtained, and the matching point pairs of the labels can be solved to obtain the monoresponse matrix between the actual ground area and the camera imaging plane.

Secondly, the delineation of a two-dimensional planar hazardous area is carried out given three-dimensional coordinate points. A quadrilateral region is uniquely determined by giving the boundary points of the desired delineated region. The endpoints B of the quadrilateral, which are occluded by the environment, can be calculated directly by the one-shoulder transformation, or their 2D pixel coordinates can be obtained by finding the intersection point of a straight line AE and a straight line CF based on the linearity preserving nature of the one-shoulder transformation.

Finally, after delineating the hazardous area, determining whether the operator has entered the area also requires determining the coordinates of the intersection position of the pedestrian target and the actual ground in the 2D imaging plane, assuming that the midpoint of the bottom of the target frame of the operator's target P is used as its current coordinate point in the 2D pixel plane.

Then the hazardous area misentry detection problem is transformed into the problem of determining whether a point in the 2D plane is inside a polygon. Generally speaking, the hazardous area delineated on the actual ground is rectangular in shape, which is projected to the 2D pixel plane as a convex quadrilateral shape, so the convex quadrilateral is used as an example to analyze and solve the problem.

There are geometric laws as follows: for a point in the plane, in clockwise direction, if the point is always located to the right of each side vector, then the point is located inside this convex quadrilateral. The parallelogram and its interior and exterior points are shown schematically in Figure 4. A point P inside the quadrilateral $ABCD$ lies to the right of the vectors $\overrightarrow{AB}$, $\overrightarrow{BC}$, $\overrightarrow{CD}$, and $\overrightarrow{DA}$, and a point Q outside the quadrilateral $ABCD$ lies to the right of the vectors $\overrightarrow{AB}$, $\overrightarrow{CD}$, and $\overrightarrow{DA}$ and to the left of the vector $\overrightarrow{BC}$. To determine whether a point lies on the right side of a vector is simply a matter of determining the crossover relationship between the two vectors. For example, the vector $\overrightarrow{AP}$ lies in the clockwise direction of the vector $\overrightarrow{AB}$, and therefore the point P lies on the right side of the vector $\overrightarrow{AB}$.

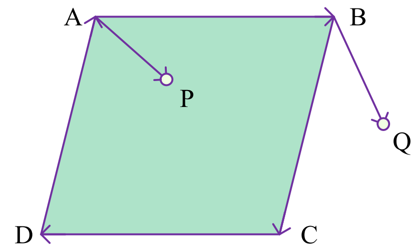

Figure 4: The illustration of a parallelogram and its internal and external points

Alternatively, the determination of the cis-inverse relationship can be accomplished by virtue of computing the vector product. Suppose two vectors $\vec{M}$, $\vec{N}$ in the plane, if the result of their vector product is 0, i.e., $\vec{M} \times \vec{N} = 0$, the vector $\vec{M}$ is covariant with the vector $\vec{N}$; if the result of their vector product is positive, i.e., $\vec{M} \times \vec{N} > 0$, then the vector $\vec{M}$ is in the clockwise direction of the vector $\vec{N}$; if the result of their vector product is negative, i.e., $\vec{M} \times \vec{N} < 0$, then the vector $\vec{M}$ is in the counterclockwise direction of the vector $\vec{N}$.

Based on the above method, it is possible to determine whether a point in the plane lies inside a convex quadrilateral. Assuming that in the 2D imaging plane, the delineated hazardous area is a convex quadrilateral $ABCD$, and the detected pedestrian target coordinate point is P, if it satisfies:

$$\begin{cases} \overrightarrow{AP} \times \overrightarrow{AB} > 0 \\ \overrightarrow{BP} \times \overrightarrow{BC} > 0 \\ \overrightarrow{CP} \times \overrightarrow{CD} > 0 \\ \overrightarrow{DP} \times \overrightarrow{DA} > 0 \end{cases} \tag{8}$$

Then the pedestrian target is mistakenly within the danger zone, otherwise it is judged that the pedestrian target is not within the danger zone.

Therefore, based on the above method, it is possible to determine whether the pedestrian target mistakenly enters the delineated hazardous area. The specific steps for detecting misentry into the hazardous area of electric power operation are as follows:

(1) In the three-dimensional world coordinate system, delineate the ground area that is prohibited to enter by giving the ground coordinate information of the boundary points of the rectangular area.

(2) Calculate the corresponding points of the boundary points of the dangerous area in the two-dimensional imaging plane through the single responsive plane projection, and project the dangerous area from the actual ground plane to the two-dimensional imaging plane of the camera.

(3) Based on the detection and tracking results of the pedestrian target, take the midpoint of the bottom of its target frame as the corresponding point of its actual ground coordinates, and combine it with the dangerous area of the 2D imaging plane to determine whether the pedestrian target has mistakenly entered the dangerous area. If the pedestrian target mistakenly enters the dangerous area, an alarm message is sent.

III. Experimentation and performance verification of substation operation safety monitoring model based on improved YOLOv5

Chapter 2 proposes a safety monitoring model for substation operation that integrates high resolution remote sensing image segmentation and improved YOLOv5, and designs a dynamic detection method for hazardous area misentry. In order to verify the effectiveness of the proposed method, Chapter 3 develops experiments from four dimensions: denoising performance, safety equipment detection, comprehensive index comparison and hazardous area misentry detection, and comprehensively evaluates the practical application capability of the model by combining quantitative analysis and visualization results.

III. A. Experimental setup

III. A. 1) Data sets

A portion of images are randomly selected from remote sensing images of a substation, including images of different feature types and different lighting conditions, and some images with noise points also need to be selected as a test dataset.

III. A. 2) Network models

The feature extraction backbone network based YOLOv5 model constructed in this paper is used as the base network, and the denoising module is added to it to realize the denoising function. The denoising module selects the multi-threshold segmentation algorithm to segment the image and remove the noise.

III. A. 3) Test environment and parameter settings

In this study, a computer with Linux operating system was used as the test platform, configured with the necessary development environment (Python 3.6, PyTorch), and the experiment was conducted using a GPU to run the model on an NVIDIA Ge-Force RTX 2080.

The specific parameters of the algorithm for this study were set as follows: detection threshold 0.5, input image size 416×416, batch size 128, learning rate 0.001, classification confidence threshold 0.6, target region size threshold 100, Otsu algorithm threshold range 2~255, and optimizer Adam.

III. A. 4) Evaluation indicators

The standard deviation (MSE), peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) are used to evaluate the denoising effect, the smaller the MSE and the larger the PSNR are, the better the denoising effect is. SSIM mainly measures the similarity of two images from three aspects: structural, luminance and contrast information, and takes the value of the interval as [-1, 1], the closer the value is to 1, it means that the two images are more similar, and vice versa. The value is [-1, 1], the closer to 1, the more similar the two images are, and vice versa.

III. B. Remote sensing image denoising effect evaluation results

After completing the experimental environment and parameter settings, this section further compares the performance of different denoising algorithms in remote sensing images of substation operations, and verifies the noise suppression ability of the combination of multi-threshold segmentation and improved YOLOv5 in this paper.

Gaussian noise and pretzel noise test: under the premise of retaining the visible features of the original image information, different degrees of SNR are set to add Gaussian noise, and the SNR interval in this study is [10, 50] dB. Secondly, under the premise of retaining the visible features of the original image information, pretzel noise is added, and the ratio of the noise to the number of pixels in the whole image is 5%.

YOLOv3, Faster RCNN, EfficientDet-D1, F3ViT, Mask RCNN-R101 and RetinaNet-R10 and the algorithms of this study were used to perform denoising experiments on remote sensing images of substation operations with Gaussian noise and pretzel noise added, respectively, and the results are shown in Table 1.

Table 1: Evaluation results of denoising effect of remote sensing images

Method	Evaluation indicators of remote sensing images containing Gaussian noise			Evaluation indicators of remote sensing images with high salt-and-pepper noise		
	PSNR/dB	MSE/dB	SSIM	PSNR/dB	MSE/dB	SSIM
YOLOv3	25.58	0.103	0.686	25.71	0.093	0.712
Faster RCNN	26.35	0.069	0.721	26.50	0.049	0.744
EfficientDet-D1	27.99	0.037	0.883	28.11	0.023	0.815
F3ViT	26.56	0.074	0.843	26.77	0.045	0.833
Mask RCNN-R101	27.77	0.055	0.835	28.05	0.034	0.792
RetinaNet-R10	28.03	0.048	0.893	28.25	0.035	0.853
OURS	28.74	0.034	0.907	28.98	0.026	0.874

The method proposed in this paper significantly outperforms other comparative models in terms of denoising effect. For images containing Gaussian noise, the PSNR value of the improved YOLOv5 target detection algorithm is 28.74dB, the MSE is 0.034, and the SSIM is 0.907; for pretzel noise images, the PSNR is 28.98dB, the MSE is 0.026, and the SSIM is 0.874, which are all optimal. In comparison, traditional models such as YOLOv3 with a PSNR value of 25.58dB and Faster RCNN with a PSNR value of 26.35dB perform weakly, while EfficientDet-D1 with a PSNR value of 27.99dB is close to but still a gap. The experimental results show that the algorithm combining multi-threshold segmentation and improved YOLOv5 in this paper is more advantageous in terms of noise suppression and detail preservation.

III. C. Research on Wearable Detection of Safety Equipment for Transformer Operations

Based on the high quality images after denoising, this section focuses on the safety equipment wearing detection of substation operators, and verifies the enhancement of the coordinate attention mechanism CA for small target detection by analyzing the recognition accuracy and model performance of the insulated suit and safety helmet.

III. C. 1) Precision comparison experiment

In order to better validate the monitoring ability of the substation operation safety monitoring algorithm designed in the article, for the data set about whether the substation employees wear insulated clothing and whether the substation employees wear safety helmets are detected, and the detection accuracy of the different models are shown in Figures 5 and 6, respectively.

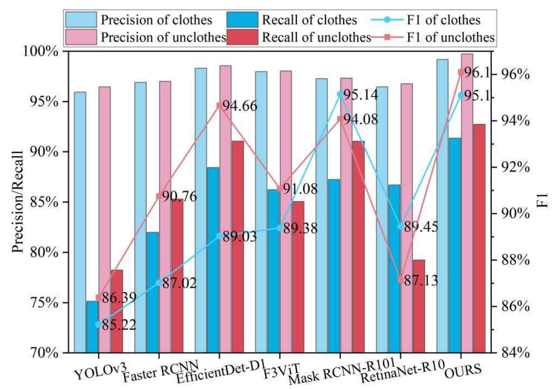

Figure 5: Accuracy of whether substation employees are wearing insulating suits

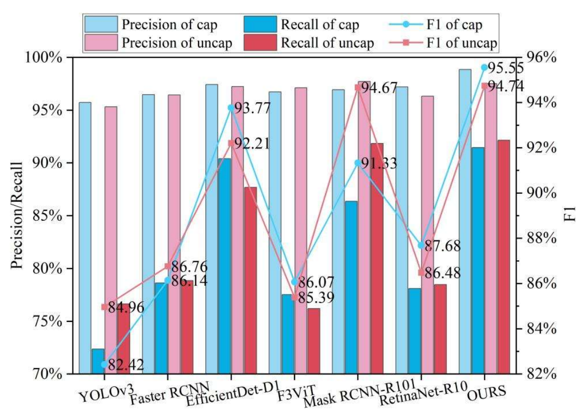

Figure 6: Accuracy of whether substation employees wear safety helmets

In insulated clothing detection, the precision rate of 99.17%, the recall rate of 91.35% and the F1 value of 95.10% of this paper's model are significantly higher than those of other models, especially the recall rate is improved by nearly 3 percentage points compared with the 88.42% of the next best EfficientDet-D1. In helmet detection, the F1 value of this paper's model reaches 95.55%, which is higher than 94.67% in Mask RCNN-R101 and 87.68% in RetinaNet-R10. The results show that the introduction of coordinate attention mechanism CA effectively improves the localization accuracy of small targets and reduces the leakage and misdetection.

III. C. 2) Performance comparison experiments

To further evaluate the comprehensive performance of the model, this section introduces the multi-scale AP value and computational efficiency index to compare the advantages and disadvantages of this paper's method with other state-of-the-art models, highlighting its balance between accuracy and efficiency.

To verify the effectiveness of the proposed method, experiments are conducted on the substation operation dataset. Continuing to compare with Mask RCNN, YOLOv5s, EfficientDet and other advanced methods, it is verified that this paper's method has a significant advantage over other methods in a number of evaluation indexes, such as AP value, detection speed and computation amount.

Table 2 shows the data of each performance index of this paper's method and other advanced methods.

Table 2: The performance index data of various methods

	AP	AP50	AP75	APS	APM	APL	Params	FLOPs
YOLOv3	32.84	57.72	34.56	18.44	35.81	41.15	59.6M	77.9B
Faster RCNN	36.29	58.18	42.84	17.56	44.92	56.81	-	-
EfficientDet-D1	39.47	58.97	43.94	18.09	45.33	57.46	8.8M	6.9B
F3ViT	35.53	60.34	36.69	19.06	43.06	46.43	6.5M	11.1B
Mask RCNN-R101	38.83	60.51	41.72	20.14	41.05	50.18	60.2M	159.6B
RetinaNet-R10	39.63	59.21	42.28	21.85	42.01	52.51	51.5M	128.4B
OURS	40.58	61.98	44.96	23.29	46.27	58.15	5.7M	7.4B

In this paper, the AP50 value of the improved YOLOv5 object detection algorithm introduced by the coordinate attention mechanism CA reaches 61.98, which exceeds the 60.51 of Mask RCNN-R101 and 60.34 of F3ViT, and the FLOPs with parameter quantity of 5.7M and computational amount of 7.4B are significantly lower than those of Mask RCNN-R101 of 60.2M and 159.6B. This indicates that the proposed model achieves a better balance between detection accuracy and computational efficiency.

The relationship between the mean average accuracy and detection speed for each method is shown in Fig. 7.

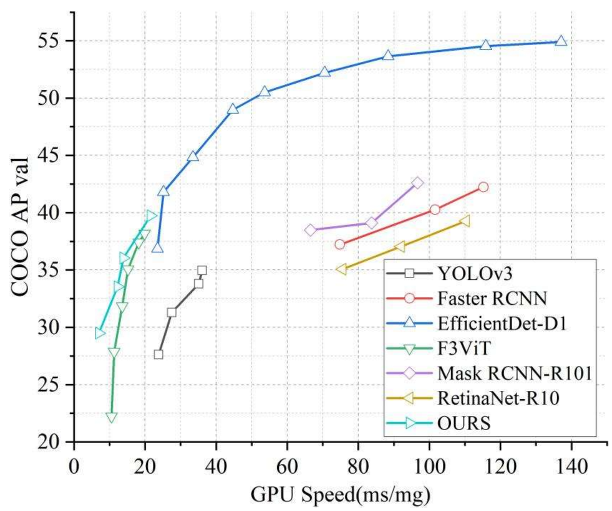

Figure 7: Relationship between average precision of each method and detection speed

Detection accuracy should be the greater the better, the detection speed is also the faster the better, so the landing point should be in the upper left of the coordinates for good, from Figure 7 can be seen, this paper's method is the most by the left on the model method, proved to have a higher accuracy at the same time, the detection speed is also faster, while the EfficientDet-D1 has a higher accuracy, but with the increase in accuracy of the detection speed is also substantially slowed down, the required detection time setting reaches 140ms/mg, while the detection speed of the model in this paper can reach 7.26ms/mg.

III. D. Research on the detection of people straying into hazardous areas

Finally, combining the single responsive transform and geometric rules, this section quantitatively analyzes the dynamic detection effect of personnel mistakenly entering dangerous areas, explores the potential causes of misdetection and omission, and improves the reliability verification of the closed loop of safety monitoring. The methods in the paper continue to be compared and analyzed with YOLOv3, Faster RCNN, EfficientDet-D1, F3ViT, Mask RCNN-R101 and RetinaNet-R10 methods.

III. D. 1) Accuracy analysis with different number of iterations

The accuracy curves of the different methods with respect to the number of iterations are shown in Fig. 8.

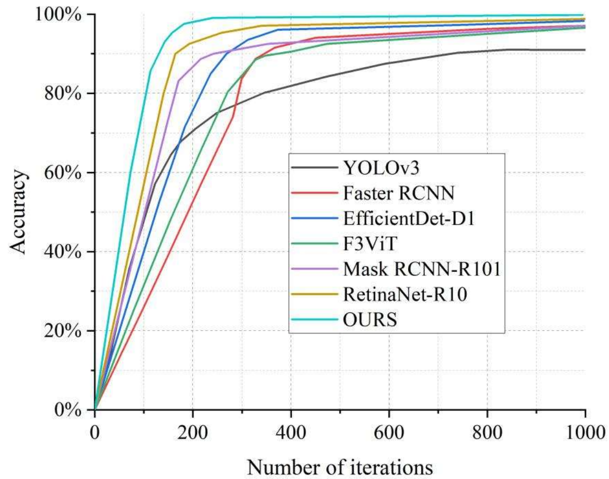

Figure 8: Accuracy of different methods varying with the number of iterations

As can be seen from Fig. 8, the accuracy of the method proposed in the paper tends to be 99.85% after about 200 iterations.The YOLOv3 method tends to be 91.02% after about 800 iterations. The Faster RCNN method converges to 96.31% accuracy after about 450 iterations. The EfficientDet-D1 method converges to 97.31% accuracy after about 400 iterations. The F3ViT method tends to be 96.58% accurate after about 500 iterations.The Mask RCNN-R101 method tends to be 97.44% accurate after about 400 iterations.The RetinaNet-R10 method tends to be 98.84% accurate after about 350 iterations.

III. D. 2) Analysis of test results

The results of the method personnel misentry hazardous area detection in the paper are shown in Table 3.

Table 3: Detection of personnel entering hazardous areas by mistake

	Dangerous area	Actual quantity	Identification quantity	Accuracy
Sample1	By mistake	168	152	90.48%
	Not entered	811	787	97.04%
	Total number	979	939	95.91%
Sample2	By mistake	67	60	89.55%
	Not entered	855	831	97.19%
	Total number	922	891	96.64%
Sample3	By mistake	107	83	77.57%
	Not entered	932	904	97.00%
	Total number	1039	987	95.00%

As can be seen from Table 3, the accuracy rate of personnel misentering the hazardous area detection is high, with three samples achieving 95.91%, 96.64% and 95.00% accuracy. And but not into the dangerous area detection accuracy rate is higher than into the dangerous area detection, may be due to the mistakenly into the dangerous area samples are less and other issues affect the accuracy of the detection. It shows that the method proposed in the paper can more accurately detect whether the substation personnel enter the dangerous area or not.

IV. Conclusion

In this study, a safety monitoring system for substation operation environment was constructed by integrating high resolution remote sensing image segmentation with the improved YOLOv5 model.

The improved model is optimal for remote sensing images containing Gaussian noise and pretzel noise, with PSNRs of 28.74 dB and 28.98 dB, MSEs as low as 0.034 and 0.026, and SSIMs of 0.907 and 0.874, which are significantly better than those of advanced models, such as YOLOv3, Faster RCNN and so on.

After the introduction of the coordinate attention mechanism CA, the detection accuracy of safety helmet and insulated clothing is significantly improved, and the F1 value reaches 95.55% and 95.10%, respectively, and the recall rate is nearly 3 percentage points higher than that of EfficientDet-D1.

The number of model parameters is 5.7M and the computational volume is 7.4B FLOPs, which is more than 90% less than Mask RCNN-R101, and the detection speed reaches 7.26 ms/mg, which takes into account high precision and real-time performance.

Hazardous area misentry detection: The dynamic detection method based on single-stress transformation and geometric rules has a combined accuracy of 95.85% in three samples, with the highest accuracy of 97.19% in the detection of un-misentered areas.

References

[1] Kitak, P., Belak, L., Pihler, J., & Ribič, J. (2021). Maintenance management of a transmission substation with optimization. Applied Sciences, 11(24), 11806.

[2] Cheng, L., Hu, W., Liu, Z., & Cai, W. (2018). On-site Smart Operation and Maintenance System for Substation Equipment Based on Mobile Network. International Journal of Online Engineering, 14(3).

[3] Yang, L., Li, W., Zhang, S., Wan, J., & Zhao, H. (2023, July). Application of Big Data and Artificial Intelligence in Substation Operation and Maintenance. In Journal of Physics: Conference Series (Vol. 2555, No. 1, p. 012002). IOP Publishing.

[4] Popa, C. (2020). Impact of substations equipment to the environment. International Journal of Global Warming, 21(2), 155-172.

[5] Dwijayatno, F. M., Bagaskara, A., Purnama, R., Kosasih, N. Z., & Kuswara, D. (2024, September). Strategic of Maintenance on Substation: Approaching with Digitalization. In 2024 International Conference on Technology and Policy in Energy and Electric Power (ICTPEP) (pp. 155-160). IEEE.

[6] Wang, J. W., Zhang, H. F., Fang, Y. G., Wen, Z. K., Shao, G. W., & Guo, D. Y. (2017). Analysis on Safety Measures of Substation Maintenance. Electronic Science Technology and Application, 4(1), 17-22.

[7] Jing, Y., Weng, L., Han, Y., Wang, X., Guo, D., & Huang, G. (2021, October). On line monitoring method of operation environment safety for substation maintenance. In 2021 International Conference on Information Control, Electrical Engineering and Rail Transit (ICEERT) (pp. 41-44). IEEE.

[8] Zhao, L., Matsuo, I. B. M., Zhou, Y., & Lee, W. J. (2019). Design of an industrial IoT-based monitoring system for power substations. IEEE Transactions on Industry Applications, 55(6), 5666-5674.

[9] Ke, Y., Chen, H., Liu, Z., & Yang, Z. (2024). Perception and early warning technology of power substation safety construction operation based on multi-dimensional information fusion. International Journal of Wireless and Mobile Computing, 26(4), 408-415.

[10] Xiong, Y., Liu, Q., Deng, M., & Niu, Z. (2021, November). Remote monitoring system design of power substation based on GPRS. In Journal of Physics: Conference Series (Vol. 2108, No. 1, p. 012031). IOP Publishing.

[11] Wang, H., Zhou, B., & Zhang, X. (2017). Research on the remote maintenance system architecture for the rapid development of smart substation in China. IEEE Transactions on Power Delivery, 33(4), 1845-1852.

[12] Zhao, B., Lan, H., Niu, Z., Zhu, H., Qian, T., & Tang, W. (2021). Detection and location of safety protective wear in power substation operation using wear-enhanced YOLOv3 algorithm. IEEE Access, 9, 125540-125549.

[13] Ullah, I., Yang, F., Khan, R., Liu, L., Yang, H., Gao, B., & Sun, K. (2017). Predictive maintenance of power substation equipment by infrared thermography using a machine-learning approach. Energies, 10(12), 1987.

[14] Wang, C., Tang, M., Hong, G., Wang, D., & Chen, Z. (2025). Dynamic integrated approach to safety risk management in substation maintenance and testing integration operations. Reliability Engineering & System Safety, 261, 111061.

[15] Wang, Y. B., Wang, Z., Zhao, D. Y., & Zhang, H. F. (2015). Intelligent Operation and Maintenance of Substations Based on Internet of Things (IoT) Technology. Applied Mechanics and Materials, 742, 708-716.

[16] Chen, W., Liu, Y., Gao, Y., Hu, J., Liao, Z., & Zhao, J. (2024). Intelligent Substation Noise Monitoring System: Design, Implementation and Evaluation. Energies, 17(13), 3083.

[17] Morais, B. T. P. (2013). Emerging technologies and future trends in substation automation systems for the protection, monitoring and control of electrical substations. PQDT-Global.