Detection and Evaluation Method of Transmission Line Defects Based on Deep Learning

HUAGANG LIANG, CHAO ZUO, AND WANGMIN WEI

College of Electronic and Control Engineering, Chang'an University, Xi'an 710064, China

Corresponding author: Huagang Liang (hgliang@chd.edu.cn)

ABSTRACT The issues of existing research on transmission line detection include the following three: only detects a few categories, no open transmission line component dataset, and no unified, comprehensive evaluation index. In this paper, we propose a detection and evaluation method of defect for transmission line inspection based on deep learning. The transmission line contains various pivotal components, while previous research has mostly focused on a few categories. In the proposed approach, the following study is performed by establishing a transmission line dataset named Wire_10, which considers defects as a category. Wire_10 contains 8 defects in transmission line components, such as insulator defect, triple-plate defect, damper defect, grading ring defect, and et al., as well as nest and foreign body that attached to the transmission line. The object detection of aerial images taken during the actual inspection is susceptible to background and lighting. These two factors are used as variables to define the background-dataset and the lighting-dataset. Faster R-CNN, an end-to-end and high recognition accuracy deep learning algorithm, is used to build detection models with transfer learning and fine-tuning. The results show that the detection method can accurately identify the defect categories in the Wire_10 dataset and is robust to aerial images with complex backgrounds and different lighting. The proposed method can effectively and accurately identify defects in the automatic inspection of transmission lines.

INDEX TERMS Transmission line defects, deep learning, aerial image, robustness.

I. INTRODUCTION

With the development of the power grid and the popularity of the unmanned aerial vehicle (UAV) inspection tour technology, the inspection of the transmission line by using the UAV has become the primary way of the defect detection of the power supply system. Instead of manual inspection, UAV inspection reduces the cost and risk of transmission line inspection. However, the traditional artificial screening and check of aerial images collected by UAV have the disadvantages of low efficiency and high missing rate of inspection. It is imperative to carry out efficient and reliable intelligent detection on aerial images. The object detection of aerial images has the problems of small proportion, complex background, multiple interferences, large variations in the contrast between object and background under different illumination conditions. Meanwhile, transmission line

The associate editor coordinating the review of this manuscript and approving it for publication was Canbing Li.

components have polysemy. For example, insulators include needle insulators, rod insulators, and suspension insulators. These problems are the difficulties in defect detection of transmission line components in aerial images.

In traditional methods, various hand-crafted features are commonly adopted. Zhu *et al.*[1] and Ban *et al.* [2] proposed the semantic model of local contour features through lines and curves to detect the spacing rod, the damper, and the grading ring. Zhang *et al.* [3] adopted Hough transform [4] and statistical texture features to form feature sequence curves for defect detection of insulators. Wang *et al.* [5] identified glass insulators and composite insulators by combining shape, color, and texture features, and diagnosed the glass insulator's off-chip defect. The algorithm for detecting objects by manual features has high customizability, and relatively, its versatility is poor. Moreover, aerial images are susceptible to cluttered environments and changes in lighting. The accuracy and reliability of such detection methods in practical applications are limited. On the other hand, the machine learning
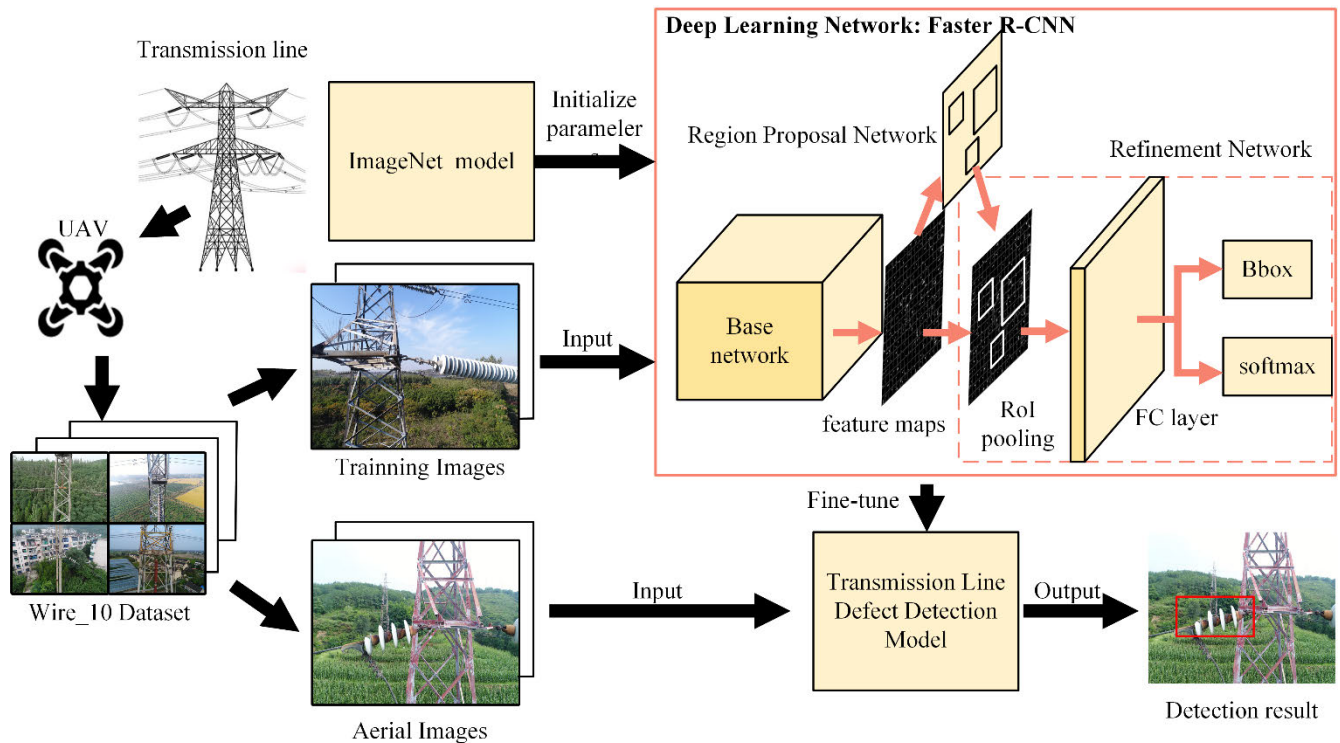
FIGURE 1. Schematic diagram of the deep learning based detection approach for defect of the transmission line. ImageNet: a large-scale hierarchical image database.

algorithm is also applied in transmission line component detection. Wu [6] proposed a method to detect insulators based on the region of interest and AdaBoost [7] cascade classifier. Zhai *et al.* [8] used the object recommendation algorithm and the structure search method, and the classifier is trained by the extracted insulator local features. The above methods have more robust feature representation capabilities and improve detection accuracy. However, they cannot handle massive high-pixel images with the shallow architectures.

Deep learning, a branch of machine learning to solve the problem of feature extraction, can automatically learn the association between features and tasks, as well as process large and complex data. Convolution Neural Network (CNN) [9] in deep learning can learn more robust and expressive features that have been applied to transmission line detection and achieved remarkable performance. Zhao *et al.* [10] utilized CNN to represent the feature of the infrared image, so as to classify insulators. Tang *et al.* [11] used Faster R-CNN [12] to detect five types of components and defects, among which the recognition accuracy of the grading ring reaches 96.8%. However, it had fewer categories and did not consider the detection effect of the algorithm under extreme conditions. Miao *et al.* [13] used single shot multibox detector (SSD) [14] to detect insulators, achieving an excellent real-time effect. Zhang [15] used the YOLOv3 [16] to detect the abnormal target in the transmission line. Although SSD and YOLO [17] improve the speed of transmission line component

of object detection, Faster R-CNN has a definite advantage in detection accuracy.

In this paper, to address the above challenges, we attempt to establish an evaluation criterion for object detection on transmission lines and propose a deep learning based method to detect transmission line defects. We build the detection model based on Faster R-CNN to detect the 10 categories, including 8 component defect types, nest, and foreign body. The transfer learning method is used to train the detection model. By comparing the recognition accuracy under different base networks and training iterations, we obtain a detection model with the best recognition effect. Then the aerial images with complex background and diverse light intensity are studied to verify the robustness of our detection method.

The rest of this paper is organized as follows. Section II describes the establishment of datasets and the proposed framework of the detection approach for transmission line defects. The experimental results on the different tasks are presented in Section III. In Section IV, we conclude from the results and also discuss our future work plans.

II. METHODOLOGY
A. FRAMEWORK OF TRANSMISSION LINE DETECTION METHOD
Fig. 1 is the schematic diagram of the detection method for transmission line inspection. The process of aerial image detection based on deep learning mainly includes
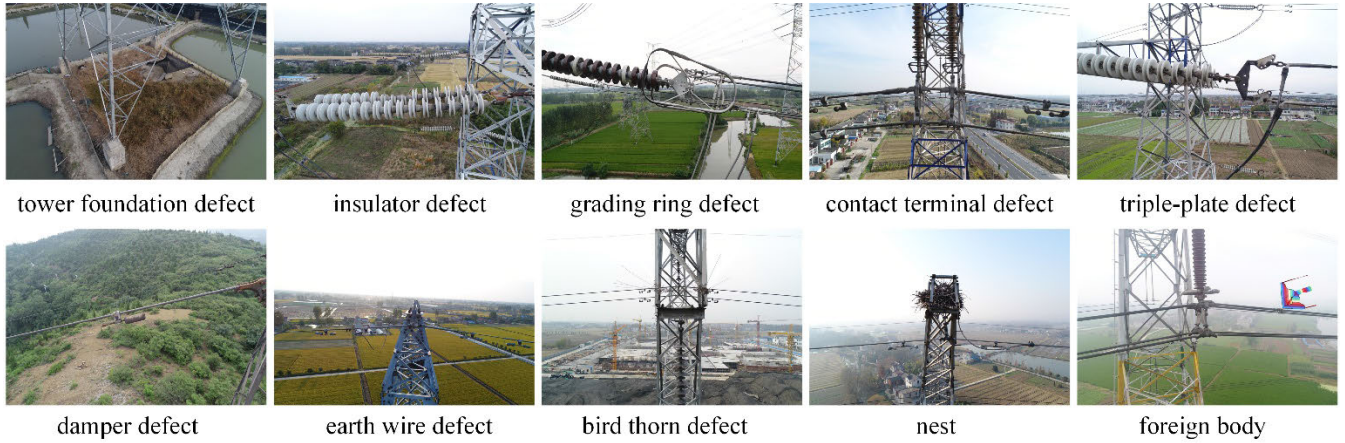
FIGURE 2. Wire_10 dataset. It includes 10 defect types.

image preprocessing, training of deep learning network Faster R-CNN, and detection of the aerial images. First, we use UAV to inspect the transmission lines, obtain aerial images, and establish a transmission line defect dataset named Wire_10 containing 10 defect types. Defining two sub-datasets: background-test dataset and light-test dataset, which are based on the background and lighting of the image, to verify the robustness of the detection model to the background and light. For the detection model, we utilize the Faster R-CNN algorithm to train by feeding the images and initialize it by ImageNet [18] model. Once the detection model of transmission line defects is well trained, it can be used to detect defects of the transmission line in aerial images taken by UAV.

B. IMAGE PROCESSING

1) WIRE_10 DATASET

At present, there is no standard and open image database in the field of transmission line components taken by UAV. Compared with the detection of transmission line components, the detection of the component defects can better meet the needs of practical applications and contribute to the inspection work. Therefore, the defects of the transmission line are treated as a category when building the dataset. In the paper, the self-built transmission line defects dataset was named Wire_10, and related research was carried out on this basis.

This dataset contains 10 categories of transmission line defects. There are 8 types of component defects: tower foundation defect, insulator defect, grading ring defect, contact terminal defect, triple-plate defect, damper defect, earth wire defect, and bird thorn defect, and 2 types of other defects: nest and foreign body. Table 1 is a detailed description of the defects contained in the Wire_10 dataset. The sample of the Wire_10 dataset is shown in Fig. 2. The descriptions in Table 1 corresponding to the component defects in Fig. 2 are as follows: the tower foundation defect is

TABLE 1. Description of defects contained in the Wire_10 dataset.

Defect Name	Defect Description
Tower foundation defect	Crop planting, stacking debris, overgrown weeds
Insulator defect	Breakage
Grading ring defect	Breakage, deformation
Contact terminal defect	Breakage, deformation
Triple-plate defect	Breakage, deformation
Damper defect	Breakage, deformation
Earth wire defect	Break
Bird thorn defect	Lack of thorns, deformation
Nest	Nest
Foreign body	Hanging foreign body

overgrown weeds, the insulator defect is breakage, the grading ring defect is deformation, the contact terminal defect is deformation, the triple-plate defect is breakage, the damper defect is breakage, the earth wire defect is break, and the bird thorn defect is lack of thorns. It should be noted that each component defect type is a general term and does not distinguish its specific genre, which is for facilitating the detection.

2) IMAGE UNDER SPECIAL CONDITIONS

During the inspection of the UAV, the background of the aerial image is very complicated due to the diverse environments. Considering the impact of different image backgrounds on recognition accuracy, this paper divides the aerial images with the river, forest, building, and farmland as the main background into different datasets, respectively. Samples of the background-dataset are shown in Fig. 3.

Similarly, the aerial image is subjected to different illumination intensities due to time and weather, and the contrast between the object and the background will also be
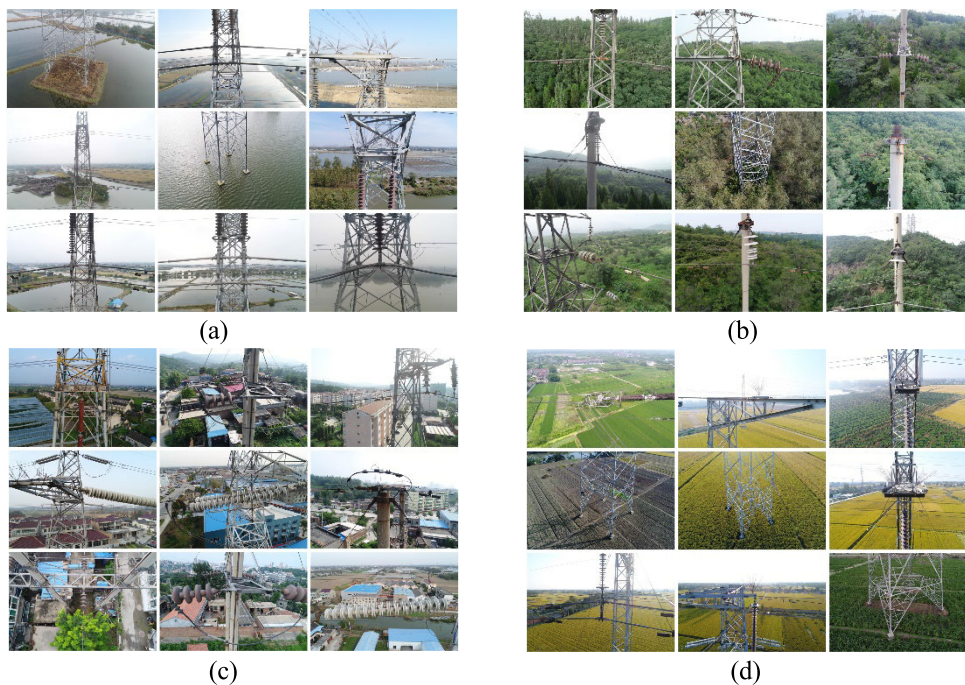
FIGURE 3. Samples of background-dataset. Aerial images with main backgrounds are (a) river, (b) forest, (c) building, (d) farmland.
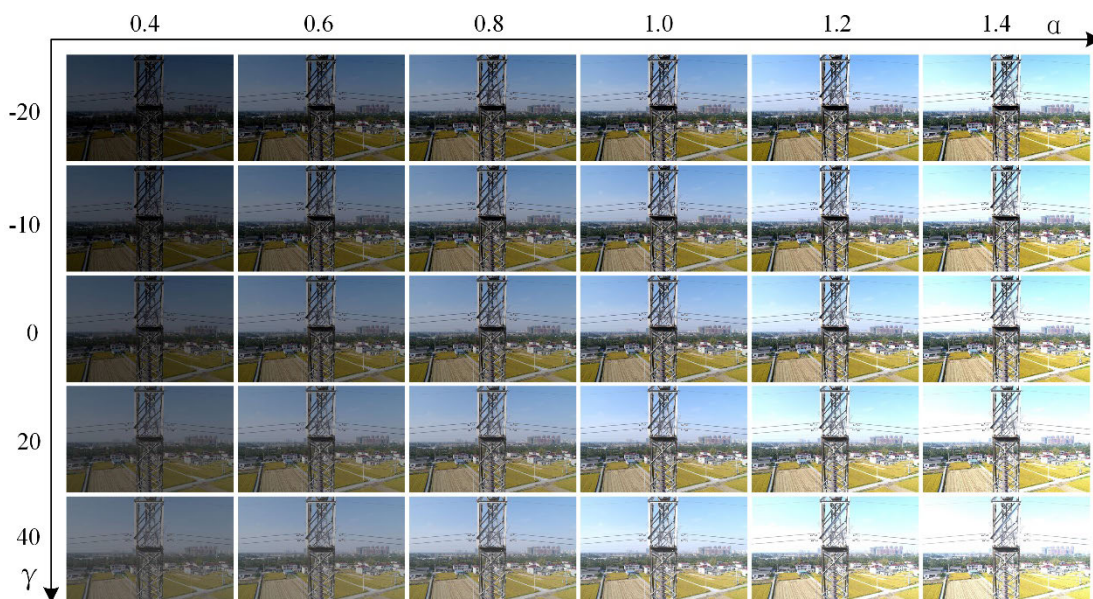

FIGURE 4. Samples of light-dataset. Aerial images with different α and γ.

changed, thus affecting the detection accuracy. We simulate the change in light intensity by adjusting the contrast and brightness of the image. Selecting the image with moderate illumination intensity as the original image, and merging the zero-pixel image with the original image to adjust the contrast and brightness. The method is as shown in the formulas (1) and (2). Fig. 4 shows images obtained by simulated illumination.

$$g_n = \alpha f_n + \beta f_0 + \gamma \qquad (1)$$

$$\beta = 1 - \alpha \qquad (2)$$

f_n is the original image, g_n is the adjusted image, and f_0 is the zero-pixel image with the same size as f_n, α is a multiple of the original image, β is a multiple of the zero-pixel image, γ is the added pixel value. The larger the α, the higher the contrast. The larger the γ, the brighter the image.

The image is dark and blurred when α is 0.4. If it continues to reduce, it will seriously affect the detection effect. In actual operation, it is prohibited to conduct UAV
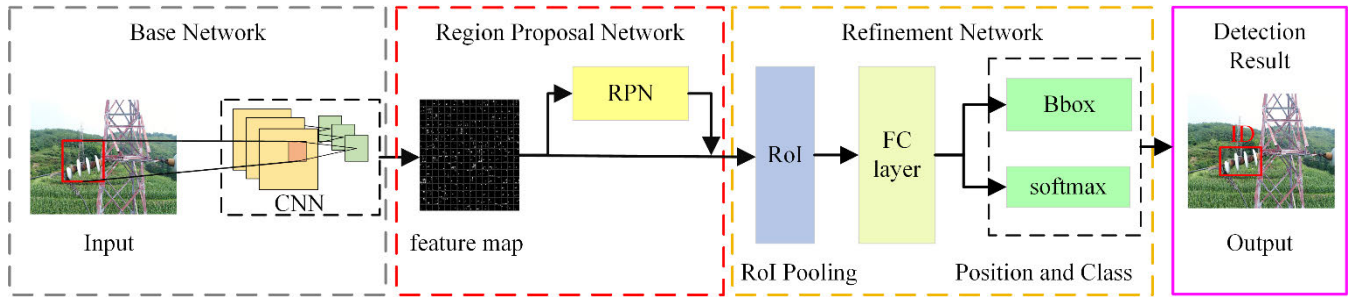
inspections with extremely low visibility. Therefore, for image contrast, α includes six different levels of 0.4, 0.6, 0.8, 1.0, 1.2, and 1.4. Similarly, the image is generally dark when γ is -20, so for brightness, γ is designed as -20, -10, 0, 20, and 40. Based on the original image with γ as 0 and α as 1, the images with γ less than 0 or α less than 1 are simulated cloudy or afternoon light; the images with γ greater than 0 or α greater than 1 are simulated brilliant or early morning light.

C. TRANSMISSION LINE DETECTION NETWORK

Faster R-CNN is a classical method of object detection that has high recognition accuracy. It is suitable for the detection task of the aerial image. The schematic diagram of the Faster R-CNN object detection network for transmission lines is illustrated in Fig. 5. The network is mainly composed of the base network, the region proposal network (RPN) [12], and the refinement network.

1) BASE NETWORK

The base network is an indispensable part of computer vision tasks, which extracts image features for subsequent tasks. The classification network with excellent performance is usually used as the base network for complex tasks such as detection and segmentation. The features extracted from these networks are more comprehensive than traditional hand-craft features. The ZF [19] and VGG-16[20] are chosen as the base network in [12]. In this paper, the base network of Faster R-CNN is VGG16, which is due to its better feature extraction ability. The VGG16 consists of 13 convolutional layers, 5 pooling layers, and 3 fully connected layers. The features of the image are extracted using the first 13 convolutional layers of VGG16. ResNet-101[21] is proposed after VGG, which adopts residual learning ideas and solves the feature loss problem to gain accuracy. Therefore, another base network of Faster R-CNN is ResNet-101 (except the fully connected layer).

The VGG16 and ResNet-101 networks differ in the strategy, depth, and complexity of extracting features, resulting in different feature maps extracted from the same image. The VGG16 and the Resnet-101 will be selected as the base networks of the detection model for transmission line defects, respectively. The recognition accuracy of different defects is studied in Section III.

2) REGION PROPOSAL NETWORKS

The Region Proposal Network is a fully convolutional network which inputs image with any size and outputs a set of rectangular object proposals. In detail, the RPN takes feature maps output by the last convolutional layer as input, then slides and scans on the feature maps with a sliding window of 3∗3 size, and finally generates a feature vector that passes to the class decision layer and the position regression layer. The class decision layer determines whether the object belongs to the foreground or the background through the probability value of the foreground and background of the proposed region of each position. The position regression layer is to locate the object coarsely. The schematic diagram of the RPN is shown in Fig. 6.

We generate 9 proposed regions (named anchors) for each sliding window in the feature map. Anchors are a set of fixed-size proposed regions with three dimensions {128^2, 256^2, 512^2} and three aspect ratios {1:1,1:2,2:1}. Then each sliding window is mapped to a low-dimensional feature (512-*d* for VGG16 and ResNet-101) and passes the low dimensional vector to the class decision layer and the position regression layer.

When training the RPN, each anchor needs to be divided into positive and negative samples, i.e., foreground and background. Therefore, the number of convolution kernel channels of the classification layer is 18, while the regression layer is 36. The position information of each anchor is described by the coordinates: x, y, and the length and width values: w, h.

The loss function includes two parts: the position regression loss function and the classification loss function. The loss function is shown in (3).

$$L(\{p_i\}, \{t_i\}) = \frac{1}{N_{clc}}\sum_i L_{clc}(p_i, p_i^*) + \lambda\frac{1}{N_{reg}}\sum_i p_i^* L_{reg}(t_i, t_i^*)$$

(3)

Here, i is the index of the anchor and p_i is the predicted probability of anchor i that being an object. The value of p_i^* is 1 if the anchor is positive, and is 0 if the anchor is negative. t_i contains the location information of the prediction proposed region, and t_i^* contains the location information of the positive anchor. L_{cls} is the classification loss function and L_{reg} is the position regression loss function. The equations of L_{cls} and L_{reg} are shown in (4), (5). R in (5) is the robust loss
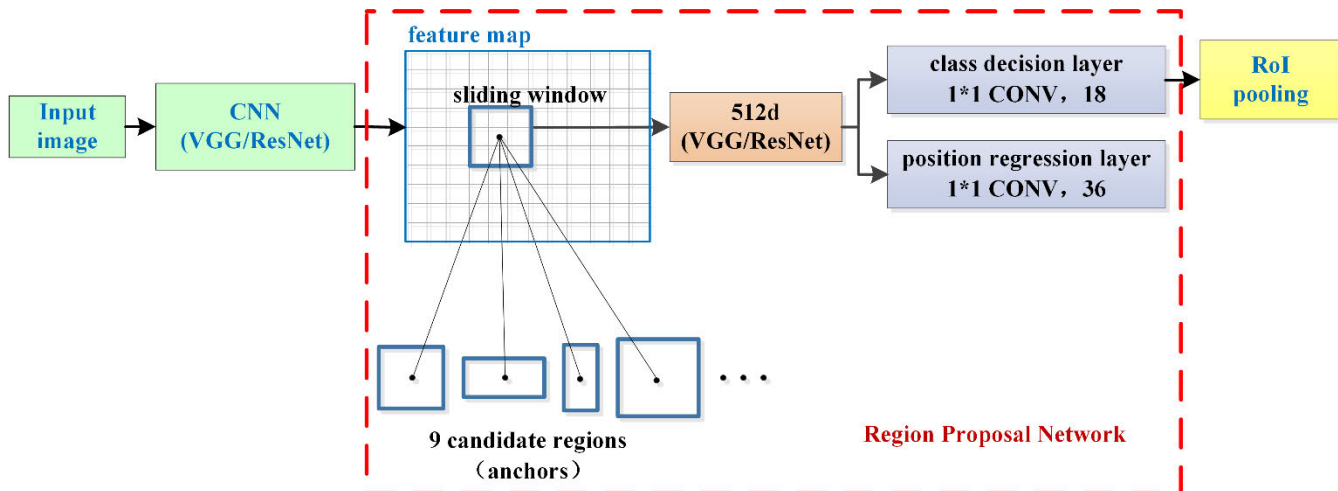
FIGURE 6. Schematic diagram of the regional proposal network.

function defined in (6).

$$L_{cls} = -\log\left[p_i p_i^* + (1 - p_i)(1 - p_i^*)\right] \quad (4)$$

$$L_{reg}(t_i, t_i^*) = R(t - t^*) \quad (5)$$

$$smooth_{L_1}(x) = \begin{cases} 0.5x^2 & |x| < 1 \\ |x| - 0.5 & \text{else} \end{cases} \quad (6)$$

3) REFINEMENT NETWORK
The refinement network for classification and position consists of the RoI pooling layer, fully connected layer, class decision layer (softmax), and position regression layer (Bounding box regression). The inputs of the network are the outputs of the base network and the region proposal network. The RoI pooling layer uses the max-pooling to normalize the fixed features map. After the fully connected layer, the category probability and the position information of the object are output through the class decision layer and the position regression layer. The refinement network structure is shown in Fig. 7.

The class decision layer determines the category of the RoI using the softmax classifier. When there are k categories, an array of $k + 1$ dimensions (k categories and backgrounds) is output by softmax, which means the probability that the object belongs to each category. In this paper, we choose the highest probability as the detection result. The probability score of each ROI is $P = (p_0, p_1 \cdots p_{10})$. For a class u, its loss function is shown in (7).

$$L_{cls}(p, u) = -\log p_u \quad (7)$$

The position regression layer uses bounding box regression to adjust location information. The output of this layer is an array of size $4*k$ dimensions indicating the value that should be panned and scaled if the object belongs to a particular category. Then determining the location of the output box. For a class u, $t^u = (t_x^u, t_y^u, t_w^u, t_h^u)$ represents the four parameters that should be translated and scaled between the ground truth and
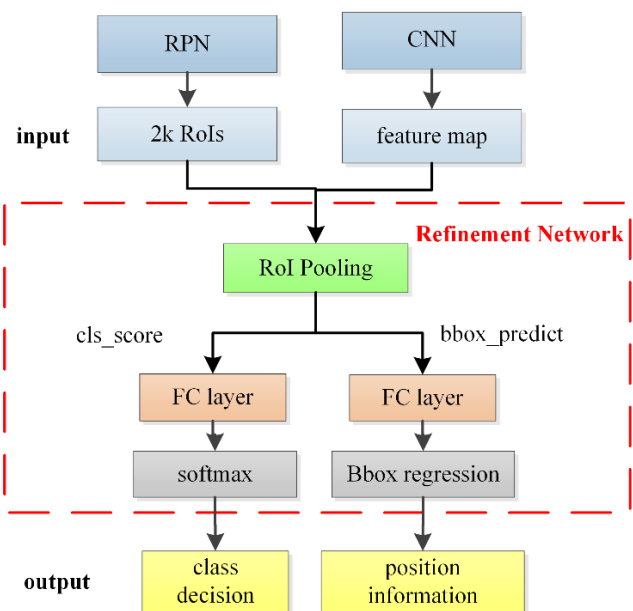
FIGURE 7. The refinement network for classification and position.

the predicted value. $v = (v_x, v_y, v_w, v_h)$ represents the ground truth labeled in the image. $t^u = (t_x^u, t_y^u, t_w^u, t_h^u)$ represents predicted value. The loss function of the position regression is shown in (8).

$$L_{loc}(t^u, v) = \sum_{i=1}^{k} smooth_{L_1}(t_i^u - v_i) \quad (8)$$

When training the refinement network, its loss function consists of the class decision loss function and the position regression loss function. As shown in (9). $L_{cls}(p, u)$ is the class decision loss function and $L_{loc}(t^u, v)$ is the position regression loss function. λ is hyper-parameter and $\lambda = 1$ in

FIGURE 8. Training flowchart of transmission line detection model.

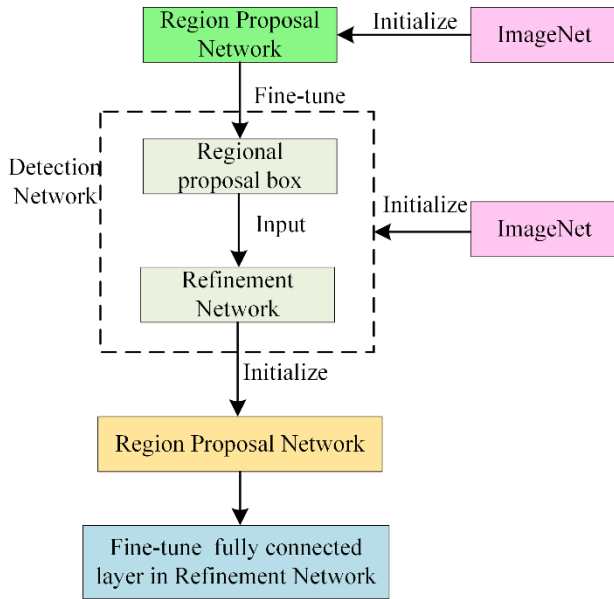the experiment.

$$L(p, u, t^u, v) = \begin{cases} L_{cls}(p, u) & u \ as \ background \\ L_{cls}(p, u) + \lambda L_{loc}(t^u, v) & else \end{cases}$$

$$(9)$$

D. MODEL TRAINING

The proposed detection network adopts the multi-task learning method by sharing convolutional layers. To ensure that the RPN and the refinement network realize the convolutional layer sharing of the base network during the training process, the alternating training and transfer learning method is used for network training. The training process is shown in Fig. 8.

The alternating training steps are as follows: 1) Initializing the region proposal network with the pre-trained ImageNet model and fine-tuning the region proposal network to get the region proposal box. 2) The ImageNet model is used again to initialize the detection network, and the regional proposal box is used as the input of the refinement network to generate an independent detection network. 3) Using the independent detection network to initialize the regional proposal network, and train after fine-tuning it. 4) Finally, the shared convolutional layer is fixed, and only the fully connected layer in the refinement network is fine-tuned so that two networks can share the convolution layers. We train the detection model of transmission line defects on the Wire_10 dataset. The recognition accuracy and generalization ability of the model will be tested in Section III.

III. EXPERIMENTAL RESULTS

In this section, we introduce the details of the evaluation, including the selection of the base network, settings of the experiments, and obtained results.

TABLE 2. Description of datasets.

Dataset	Data size	Note
Wire_10 dataset	8000	/
Background-dataset	800	4 backgrounds
Light-dataset	6000	30 lighting conditions

A. DATA PREPARATION

We conduct all experiments on the self-built dataset: Wire_10. The aerial images are provided by China Southern Power Grid and Yangquan Coal Industry. They were collected by the "Tian Xuan II" UAV on the power transmission line. The shooting time includes different time points in summer and winter. The original image size is $5472 * 3648$. As shown in Table 2, there are 3 datasets used in the experiments. The detection network learns the detection model of transmission line defects by feeding the Wire_10 dataset and verifies the model's robustness to the background and lighting with the background-dataset and the light-dataset. The Wire_10 dataset contains 8000 images, including 10 defect types acquired during inspections of different transmission lines. The images include cluttered backgrounds and various lighting conditions. The Wire_10 dataset has a ratio of 7: 3 for training and testing. In the background-dataset, there are 200 images for each of 4 backgrounds. The number of images for each defect category on a particular background is 20. In the light-dataset, there are 30 lighting conditions divided by adjusting the contrast and brightness. There are 200 images for each lighting condition, and each category has 20 images in one lighting condition. The images are labeled with the deep learning annotation tool labelImg. For the transmission line defects that appear totally in the image, we mark its outer frame and category label, and each image marks one object.

B. IMPLEMENTATION

The proposed approach is implemented in Caffe [22]. It is running on a computer equipped with Intel(R) Core (TM)i7-7700K CPU, an NVIDIA GeForce GTX 1080 TI GPU, and 16GB of RAM under Ubuntu16.04.

The parameters of the proposed method are set as follows. The initial learning rate is 0.001, the decay factor is 0.1, and the momentum is 0.9. L2 regularization is used to avoid overfitting problems in the model. The base network selects VGG16 and ResNet-101, respectively. There are two options for training iterations, 240,000 and 400,000. Since aerial images are high-definition images, CNN will inevitably suffer from insufficient memory due to excessive calculations. Therefore, the shortest edge scale of the image is set to 300, the number of regions of interest is 128, and the maximum batch size is 128. During the actual inspection, the false detection and leakage detection have a powerful influence on the inspection work. We applied recall, precision, false alarm, average precision (AP), and mean average precision (mAP) to evaluate the detection performance.

TABLE 3. Detection results of transmission line defects. Defect type abbreviated as the first letter of the word.

Detection model	Base network	Iterations /million times	GRD	DD	ID	TPD	CTD	BTD	TFD	EWD	N	FB	mAP (%)
Faster R-CNN A	VGG16	24	96.56	66.86	74.75	85.17	74.75	68.31	94.80	73.24	72.91	90.91	79.83
Faster R-CNN B	VGG16	40	98.44	74.12	78.65	92.43	77.64	75.42	95.12	79.24	75.86	94.06	84.10
Faster R-CNN C	ResNet-101	24	98.68	78.41	86.67	96.86	81.67	78.21	95.27	80.43	83.32	97.34	87.69
Faster R-CNN D	ResNet-101	40	99.72	83.23	89.56	98.08	85.79	84.62	98.28	84.11	88.87	98.86	**91.11**

tower foundation defect   insulator defect   grading ring defect   contact terminal defect   triple-plate defect

damper defect   earth wire defect   bird thorn defect   nest   foreign body
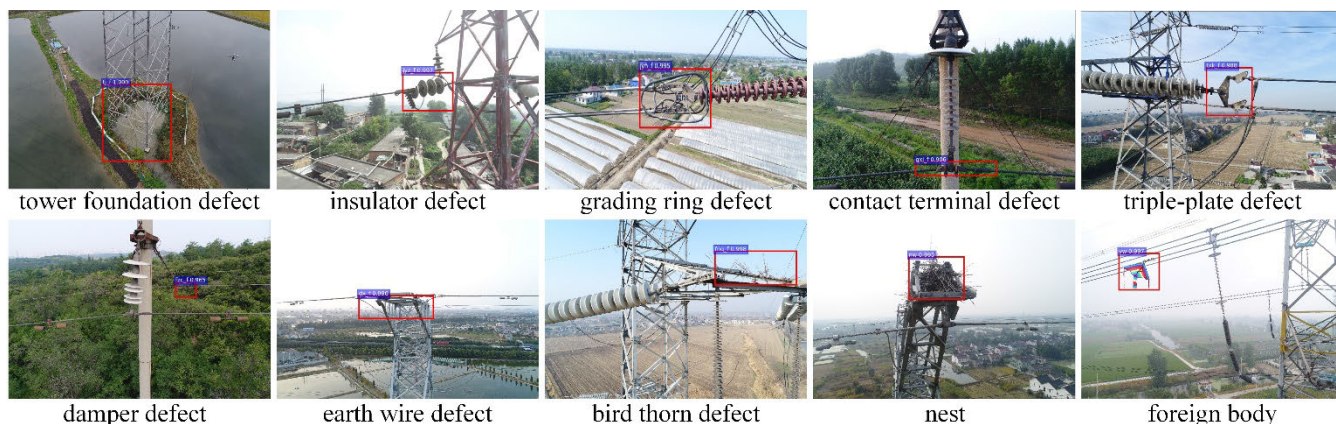
FIGURE 9. Display of some recognition results of the aerial images.

TABLE 4. Results of model comparison.

Base network	Iterations /million times	Recall (%)	Precision (%)	False Alarm (%)
VGG16	24	75.13	96.68	2.58
VGG16	40	80.29	97.47	2.08
ResNet-101	24	85.38	98.75	1.08
ResNet-101	40	88.50	99.25	**0.68**

C. DETECTION RESULTS

Our framework uses VGG16 and ResNet-101 as the base network and trains for 24 and 40 million iterations, respectively. The experimental results for 10 defect types are shown in Table 3.

According to the detection result of the transmission line defects, the mean average precision (mAP) of the base network ResNet-101 is higher than that of the VGG16 under the same number of iterations. At the same time, the mAP of the network is different under different iteration times. Under the same base network, the more iterations, the higher the accuracy of the model. The detection model Faster R-CNN D has the highest accuracy for each defect, and the AP of grading ring defect reaches 99.72%, the triple-plate defect, tower foundation defect, and foreign body defect are more than 98%, the average accuracy reached 91.1%.

Table 4 shows the other results of the four models. Faster R-CNN D has the highest recall and precision, and the false alarm reaches 0.68. Therefore, Faster R-CNN D is selected as

the follow-up research detection model. The part of the recognition results is shown in Fig. 9. It can be seen from Fig. 9 that the model has a better recognition effect on transmission line defects.

D. ROBUST EXPERIMENT

In practical applications, the aerial images collected by UAV inspections often have complex and diverse backgrounds. The backgrounds of images in different seasons and geographic environments vary greatly. At the same time, due to the impact of shooting time, the light intensity of the images changes acutely. Therefore, this section will verify the robustness and generalization of the trained detection model for transmission line defects.

1) EFFECT OF COMPLEX BACKGROUND

This part uses the background-dataset to analyze the detection performance of the proposed detection model for defect types in different backgrounds. There are four main backgrounds: building, forest, river, and farmland. The experimental results are shown in Table 5. The name of each category is represented by the first letter.

For different defect types, triple-plate defects have the highest recognition accuracy in the above 4 backgrounds, reaching 99.0%, and the recognition accuracy is more than 90% except for bird thorn defects and earth wire defects. For bird thorn, it may be attributed to the absence of a standard theoretical basis for non-defected bird thorn, and the defects were determined manually, i.e., it is not clear how

TABLE 5. Test results of transmission line defects under different backgrounds.

Background	GRD	DD	ID	TPD	CTD	BTD	TFD	EWD	N	FB	mAP (%)
building	100.0	95.0	95.0	95.0	90.0	85.0	100.0	85.0	100.0	90.0	93.5
farmland	100.0	100.0	100.0	100.0	95.0	90.0	100.0	90.0	100.0	100.0	**97.5**
forest	95.0	90.0	100.0	100.0	80.0	80.0	100.0	80.0	95.0	95.0	91.5
river	90.0	95.0	85.0	100.0	90.0	90.0	95.0	90.0	90.0	95.0	92.0
mAP(%)	95.0	96.0	92.0	**99.0**	90.0	88.0	96.0	88.0	97.0	96.0	**93.6**

TABLE 6. mAP under different α and γ.

| γ	α					
	0.4	0.6	0.8	1.0	1.2	1.4
-20	**61.5**	69.0	79.0	89.5	92.0	92.5
-10	69.0	76.0	88.0	93.0	92.5	94.0
0	76.0	82.5	92.5	93.0	93.5	94.0
20	76.5	82.5	92.5	93.0	94.0	92.5
40	78.5	85.0	93.0	**95.0**	94.0	92.0

many bird thorns are classified as defect of bird thorn, so it has a significant influence on the final recognition result of bird thorn defect. As for the earth wire, its color is similar to the background of the image. After the multi-layer convolution operation, the features of the earth wire are lost, resulting in a lower recognition result compared to other defects. For different background types, the mAP reaches 97.5% in images with farmland background, and the mAP is above 91.5% for all background types. For farmland, it is generally empty and has a single color, so it has less impact on object detection. However, for building, forest, and river, other irrelevant objects and various colors often appear in them, making detection difficult. So their accuracy is slightly worse than that of farmland.

In general, for the aerial images of 4 different backgrounds, the average mAP of 10 types of transmission line defect reached 93.6%. It can be considered that the detection model of transmission line defects in this paper has strong robustness under complex backgrounds.

2) EFFECT OF LIGHT INTENSITY

For the light-dataset, the contrast and brightness of the image are adjusted to simulate the lighting changes during the inspection, and the robustness of the detection model to lighting changes is verified. For different α and γ, the test results are shown in Table 6.

The detected images had the highest mAP reaches 95.0% when γ and α were 40 and 1.0. For α in the range of 0.8-1.4 and γ in the range of -10-40, the lighting conditions of the image are in line with most practical situations. The mAP of the detection model was almost unchanged, except that α was 0.8 and γ was -10, both of which were above 92.0%. When γ and α are -20 and 0.4, respectively,
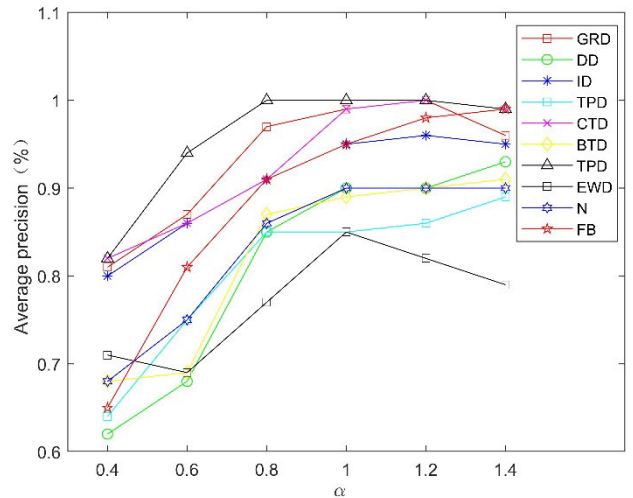
(a)
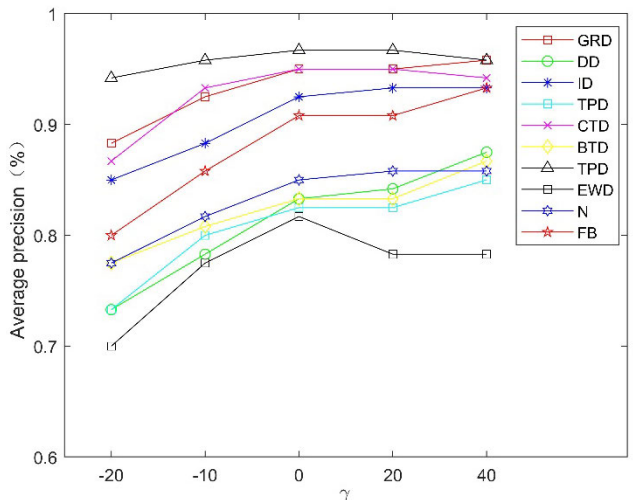
(b)

FIGURE 10. Identification results for different defect types under different conditions. (a) for different contrast and (b) for different brightness.

mAP is 61.5%. However, this case is a result obtained under extreme environments.

Then, we conduct separate experiments under different α and γ conditions to discuss the effects of contrast and brightness. The experimental results are shown in Fig. 10. Fig. 10 (a) shows the results with different α. As α increases,

TABLE 7. Comparison with other methods.

Method	Recall(%)	Precision(%)	mAP(%)	Time(ms)
Fast R-CNN+VGG16	84.69	89.24	87.25	2996.01
YOLO+ResNet101	70.47	82.88	66.33	81.40
SSD300+ResNet101	86.10	92.67.	89.65	63.12
The proposed method	88.50	99.25	91.11	288.35

the overall mAP of each defect shows an upward trend, but when α is greater than 1, the mAP of earth wire defect decreases slightly. The main reason is that there are many high voltage wires in the background of the image, and the color of the tower poles is similar to the earth wire. Fig. 10 (b) shows the experimental results obtained under different γ. With the increase of γ, the overall mAP also slowly increases. When γ is greater than 0, all mAPs are above 80% except earth wire defect. The main reason is that when the brightness of the image is very intense, the color of the earth wire is similar to the tower poles in the background, which causes the recognition accuracy to decrease.

When α and γ gradually increase, the contrast and brightness of the image increase, then the image becomes sharper. Then the features extracted by CNN will be more effective, resulting in an increase in recognition accuracy. However, excessive brightness and contrast will reduce the recognition accuracy, as some objects (e.g., earth wire) will become blurred.

On the whole, increasing the contrast and brightness of the image advisably can improve the recognition accuracy. By testing on images that simulate changes in illumination, we can draw conclusions that the proposed detection model for transmission line defects can effectively identify and detect images taken under different light intensities such as cloudy, sunny, morning, and afternoon, i.e., the detection model is robust to different light intensities.

E. COMPARISON WITH OTHER METHODS

The proposed detection approach is compared with the other three deep learning based object detection architectures: Fast R-CNN [23], YOLO, and SSD, since traditional methods were only applied for a single component category. Fast R-CNN, the predecessor of Faster R-CNN, is a region proposal based method, while YOLO and SSD are regression-based methods. They were trained in a normal way with the Wire_10 dataset. The comparison results with other methods are shown in Table 7. It shows that the proposed method has excellent recall and precision compare to the previous work. Faster R-CNN combines region proposal and CNN classification by introducing RPN to form an end-to-end detection network, which improves the accuracy and speed of detection. YOLO and SSD convert the object detection task into a regression problem, which speeds up the computation time for a single image but also reduces the accuracy. The proposed method has excellent recognition accuracy for 10 defect categories in the Wire_10 dataset.

After the model is well-trained, we can apply it to the transmission line inspection system. The defect detection method is a momentous technique and necessary prerequisite for intelligent inspection of transmission lines. During the actual inspection, the aerial image obtained by the UAV can accurately detect the defects of the transmission line through the detection model, and guide the subsequent maintenance work.

IV. CONCLUSION

In this paper, the object detection technology in computer vision is used for the detection of transmission line defects based on aerial images. Instead of the traditional manual detection, it can achieve the intelligence of image recognition and reduce the labor intensity. Aiming at the problem that there is no open and standard dataset in the field of transmission line components, a transmission line defect dataset named Wire_10 was established by UAV, and the detected categories were increased to 10, including 8 types of component defects, nest, and foreign body. Based on the region proposal network, a defect detection network for transmission line defects was constructed. The best network model was selected by comparing the recognition accuracy of different iteration times and base networks. The experimental results show that the detection model achieves the mAP of 91.1% and the false rate of 0.68% on the Wire_10 dataset. It is also verified that the model is robust to aerial images under complex backgrounds and different light intensities. This method has practical engineering application value.

The method proposed has been put into use, and the detection effect of transmission line defects can meet the practical application requirements. However, due to the limited data obtained through UAV, there are some false detections and missed detection problems in winter and non-rural areas. We will gradually expand and improve the Wire_10 dataset in future work. At the same time, the current dataset is coarse in the classification of 10 categories, and the specific classification of transmission line components is very intricate in practice. We will refine the types of various component defects and standardize the dataset further to enable transmission line inspection to obtain more accurate and useful information.

REFERENCES

[1] M. Zhu, J. Han, and D. Lu, "Object recognition method based on local contour feature in natural scenario," *Comput. Eng. Appl.*, vol. 52, no. 1, pp. 162–167, Jan. 2016.

[2] X. Ban, J. Han, and D. Lu, "Similar circular object recognition method based on local contour feature in natural scenario," *J. Comput. Appl.*, vol. 36, no. 5, pp. 1399–1403, May 2016.

[3] X. Zhang, J. An, and F. Chen, ''A method of insulator fault detection from airborne images,'' in *Proc. 2nd WRI Global Congr. Intell. Syst.*, Wuhan, China, Dec. 2010, pp. 200–203.

[4] P. V. C. Hough, ''Method and means for recognizing complex patterns,'' U.S. Patent 3069654 A, Dec. 18, 1962.

[5] W. Wang, Y. Wang, J. Han, and Y. Liu, ''Recognition and drop-off detection of insulator based on aerial image,'' in *Proc. 9th Int. Symp. Comput. Intell. Design (ISCID)*, vol. 1, Dec. 2016, pp. 162–167.

[6] Y. Wu, ''Research on insulator recognition methods in aerial image based on machine learning,'' M.S. thesis, Dept. Ctrl. Comput. Eng., NCEP Univ., Baoding, China, 2016.

[7] Y. Freund and R. E. Schapire, ''A decision-theoretic generalization of on-line learning and an application to boosting,'' *J. Comput. Syst. Sci.*, vol. 55, no. 1, pp. 119–139, Aug. 1997.

[8] Y. Zhai, D. Wang, and Z. Zhao, ''Recognition method of insulator based on object proposals and structure research,'' *J. North China Electr. Power Univ. (Natural Sci. Ed.)*, vol. 43, no. 4, pp. 66–71 and 78, Jul. 2016.

[9] A. Krizhevsky, I. Sutskever, and G. E. Hinton, ''ImageNet classification with deep convolutional neural networks,'' in *Proc. Int. Conf. Neural Inf. Process. (ICONIP)*, Doha, Qatar, 2012, pp. 1097–1105.

[10] Z. Zhao, X. Fan, G. Xu, L. Zhang, Y. Qi, and K. Zhang, ''Aggregating deep convolutional feature maps for insulator detection in infrared images,'' *IEEE Access*, vol. 5, pp. 21831–21839, 2017.

[11] Y. Tang, J. Han, and W. Wei, ''Research on part recognition and defect detection of transmission line in deep learning,'' *Electron. Meas. Technol.*, vol. 41, no. 6, pp. 60–65, Mar. 2018.

[12] S. Ren, K. He, R. Girshick, and J. Sun, ''Faster CNN: Towards real-time object detection with region proposal networks,'' *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1137–1149, Jun. 2017.

[13] X. Miao, X. Liu, J. Chen, S. Zhuang, J. Fan, and H. Jiang, ''Insulator detection in aerial images for transmission line inspection using single shot multibox detector,'' *IEEE Access*, vol. 7, pp. 9945–9956, 2019.

[14] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, ''SSD: Single shot MultiBox detector,'' in *Proc. Eur. Conf. Comput. Vis.*, 2014, pp. 740–755.

[15] X. Zhang, L. Zhang, and D. Li, ''Transmission line abnormal target detection based on machine learning yolo v3,'' in *Proc. Int. Conf. Adv. Mech. Syst. (ICAMechS)*, Kusatsu, Japan, Aug. 2019, pp. 344–348.

[16] J. Redmon and A. Farhadi, ''Yolov3: An incremental improvement,'' Apr. 2018, *arXiv:1804.02767*. [Online]. Available: https://arxiv.org/abs/1804.02767

[17] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, ''You only look once: Unified, real-time object detection,'' in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Las Vegas, NV, USA, Jun. 2016, pp. 779–788.

[18] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, ''ImageNet: A large-scale hierarchical image database,'' in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Miami, FL, USA, Jun. 2009, pp. 20–25.

[19] M. D. Zeiler and R. Fergus, ''Visualizing and understanding convolutional networks,'' in *Computer Vision—ECCV*. Zürich, Switzerland: Springer, 2014, pp. 818–833.

[20] J. Fan, T. Zhao, Z. Kuang, Y. Zheng, J. Zhang, J. Yu, and J. Peng, ''HD-MTL: Hierarchical deep multi-task learning for large-scale visual recognition,'' *IEEE Trans. Image Process.*, vol. 26, no. 4, pp. 1923–1938, Apr. 2017.

[21] K. He, X. Zhang, S. Ren, and J. Sun, ''Deep residual learning for image recognition,'' in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.

[22] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, and T. Darrell, ''Caffe: Convolutional architecture for fast feature embedding,'' in *Proc. ACM Int. Conf. Multimedia (MM)*, Jun. 2014, pp. 675–678.

[23] R. Girshick, ''Fast R-CNN,'' in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Santiago, Chile, Dec. 2015, pp. 1440–1448.

HUAGANG LIANG received the B.S. degree from the Xi'an University of Technology, China, in 2003, and the M.S. and Ph.D. degrees from Fukui University, Fukui, Japan, in 2006 and 2009, respectively. He is currently an Associate Professor with the College of Electronic and Control Engineering, Chang'an University. His research interests include pattern recognition, machine vision, and intelligent transportation.

CHAO ZUO received the B.S. degree from the North China University of Water Resources and Electric Power, Henan, China, in 2018. She is currently pursuing the M.S. degree in control science and engineering from Chang'an University. Her research interests include pattern recognition and object detection.

WANGMEI WEI received the B.S. degree from the Southwest University of Science and Technology, China, in 2015. He is currently pursuing the M.S. degree in control engineering from Chang'an University. His research interests include pattern recognition and deep learning.

● ● ●