Defect Detection Method of Overhead Line Pins Based on Multi-Sensor Data Acquisition of UAV

Xiaokaiti Maihebubai(✉)

State Grid Xinjiang Electric Power Co., Ltd., Information and Communication Company, Urumqi 830001, China
xjxjsm89@yeah.net

Abstract. Common defects such as loose and missing pins on power fittings seriously affect the stable operation of the power system. In order to find and eliminate defects in time, a method for detecting pin defects in overhead lines based on multi-sensor data acquisition by UAV is proposed. UAV multi-sensors are used to collect images of overhead line pins, and grayscale, denoising and enhancement processing are performed on the images. On this basis, a watershed based image segmentation method is proposed to eliminate background interference. Defect detection is realized based on Faster-RCNN network, features are extracted through hierarchical residual convolution module, RPN is used to delineate regions of interest, Fast-RCNN network is applied to realize category confirmation, and detection results are output. The experimental results show that under the application of the research method, the AP values of the five types of samples are all above 0.9, indicating that the research method has a high defect detection accuracy.

Keywords: UAV · Multi-Sensor · Data Acquisition · Pin Image · Preprocessing · Image Segmentation · Defect Detection

1 Introduction

As key components such as pins on high-voltage transmission towers have been exposed to wind, rain and sunlight for a long time, safety hazards such as pins falling off are prone to occur, resulting in failure of transmission lines. This kind of accident has brought great economic losses to the power company, and also brought great inconvenience to the residents' living power supply. To ensure the safety of transmission lines, it is necessary to conduct timely inspection on this key component on the high-voltage iron tower [1]. At present, the inspection of transmission lines mainly relies on manual operations, which have problems of low efficiency and poor safety, and the inspection results are greatly affected by objective factors such as personnel skills, weather, and environment. Therefore, intelligent high-voltage tower inspection is very necessary. In order to realize the automatic inspection of power transmission network based on vision, researchers have explored in multiple research directions. Helicopters, UAVs, etc., as aerial flight platforms, have the characteristics of high efficiency, accuracy and safety. Using the cameras

B. Wang et al. (Eds.): ICMTEL 2023, LNICST 532, pp. 50–64, 2024.
https://doi.org/10.1007/978-3-031-50571-3_5

installed on the flight platforms, a large number of images including key information such as insulators can be obtained. They have thus become an important way of inspecting power transmission equipment in recent years. The State Grid Corporation of China took the lead in putting drones into the inspection of high-voltage lines and collected a large number of pictures. However, after the drones transmit these pictures back to the ground, a large number of professionals are still required to look at them one by one, looking for the defective pins on the high-voltage towers from the pictures. Although this method solves the security problem, it has low efficiency and high missed detection rate, making it difficult to achieve good troubleshooting efficiency. This makes it difficult to detect the potential defects of insulators accurately, which greatly increases the maintenance cost [2]. Therefore, how to use image processing technology to automate the entire process to achieve real intelligent detection has become a research direction in this field.

The inspection of overhead line pins is usually divided into two steps: the positioning of the inspection object and the defect identification of the inspection object. Images captured by aerial inspection platforms that contain critical components often also include a variety of cluttered backgrounds such as mountains, rivers, grasslands, and farmland. In the actual detection environment, aerial images have different perspective transformations for changes and lighting conditions. Processing these images is complex and can easily lead to false detection results. Therefore, locating the inspected critical components in the image requires overcoming these disadvantages. Aiming at this point, a method for detecting pin defects in overhead lines based on multi-sensor data acquisition by UAV is proposed. The pin image is collected by UAV multi-sensor, and the collected image is preprocessed to enhance the contrast of the image and reduce the detection difficulty. The use of watershed algorithm to segment the image of overhead line pins improves the accuracy of image detection, and the Faster-RCNN network is used to complete the identification of pin defects of overhead lines.

2 UAV Multi-Sensor Capture Pin Image

In power transmission lines, pins are widely used as fasteners to connect components to stabilize the entire structure. However, due to long-term exposure to the natural environment, bolts are prone to breakage and pins are lost, resulting in hidden dangers such as large-scale transmission line failures. According to the “Administrative Regulations on Operation and Maintenance of Overhead Transmission Lines” (2016 edition), the lack of pins is a critical defect. Therefore, it is necessary to conduct regular inspections of transmission lines, and timely check and repair line equipment to ensure the safe and stable operation of the power grid. Traditional transmission line inspections mostly use manual methods, which are intensive and long-term. It is difficult to carry out line inspections in some areas with complex or even dangerous terrain. Compared with the traditional manual inspection method, the current general aerial photogrammetry technology reduces the field labor intensity of personnel to a certain extent and shortens the construction period. However, in the later stage, experienced auditors mainly check the large number of inspection pictures brought back by the drone, this not only consumes a lot of manpower and resources, but also easily leads to issues such as false positives

and missed detections [3]. In recent years, with the continuous progress of sensors and remote sensing technology, it has become possible to use multi-sensors of unmanned helicopters to obtain multi-source data, which provides new ideas and favorable conditions for solving the above problems. The use of airborne sensor system for power line inspection is a high-tech that has been widely used in recent years. Using this technology, high-resolution imagery and accurate spatial 3D information can be obtained directly in the power line corridor by flight without power failure. By using this method, the efficiency of transmission line testing can be greatly improved, and a large amount of field work can be saved, saving line testing costs. The data obtained by inspection can not only provide reference for the ledger data of transmission lines, but also provide multi-source data support for power grid management and maintenance through professional analysis and processing (such as elevation analysis, 3D visualization management, etc.).

As an important method of obtaining spatial data, unmanned aerial vehicles (UAVs) have the advantages of long endurance, low flight costs, high data resolution, and flexible scheduling for low altitude data acquisition technology. This system can achieve real-time information transmission and detection of high-risk areas. Combining with conventional aerial photography, it becomes an important supplement to satellite remote sensing and conventional aerial measurement, and prospers together with conventional aerial measurement. Especially in emergency response to major natural disasters, low altitude optical image acquisition under cloudy weather, local real-time remote sensing, and distributed daily low altitude remote sensing monitoring, unmanned aerial vehicle aerial photography systems have the advantages of both satellite remote sensing and conventional aerial photography that cannot be replaced. The use of unmanned aerial vehicles for transmission line inspection and power grid safety evaluation can leverage the advantages of unmanned aerial vehicles in terms of low cost and high timeliness, and overcome many problems faced by existing manual methods [4].

The UAV multi-sensor data acquisition system used in this paper can acquire massive high-precision airborne laser scanning (LiDAR) point cloud data, high-resolution aerial digital images, thermal infrared images, and ultraviolet images of power line corridors. The data acquisition system is mainly composed of the following parts: laser, visible light detector, infrared thermal imager, ultraviolet camera, receiver, global positioning system (GPS), inertial measurement unit (IMU)) composed of positioning and orientation system (position and orientation system, POS) [5]. The specific implementation of the multi-sensor capture pin image of the UAV is as follows:

Firstly, input the GPS timing signal and PPS into the time synchronization controller to achieve time synchronization of the time synchronization controller. The synchronous control device first provides time for laser scanners, infrared thermal imagers, ultraviolet cameras, and visible light cameras. After each sensor receives the PPS signal, reset the secondary counter to zero, which can achieve microsecond level measurement accuracy. At the same time, the synchronous control device can also activate the exposure of the camera and transmit the exposure time to the computer for recording. In this way, the time of data such as POS, visible light cameras, laser scanners, infrared thermal imagers, and ultraviolet cameras becomes a GPS time, achieving the goal of multiple data sources referencing at the same time. On this basis, the high-frequency (100–200 Hz) data obtained through POS is used to obtain the spatial pose information corresponding to the time

of each frame of data obtained by each sensor, achieving the unity of spatial coordinates. Different from the traditional multi-sensor independent control data acquisition method, the multi-sensor synchronous control method proposed in this paper can realize the unification of time and space reference of multi-sensor data acquisition. The collected data frames of laser point cloud data, optical image data, infrared video data and ultraviolet image data are synchronous observation data, which lays a data foundation for multi-source data comparison and analysis to realize fault diagnosis.

The UAV multi-sensor acquisition pin image flow chart is shown in Fig. 1.

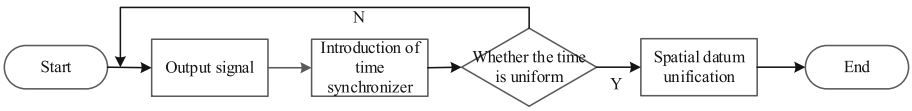


Fig. 1. UAV multi-sensor acquisition pin image flow chart

3 Overhead Line Pin Image Preprocessing

3.1 Grayscale

The grayscale of the image is a picture in which the value of each pixel on the original color image is set to R = G = B. Among them, each value of R, G, and B ranges from 0 to 255, in this way, when storing data, only one set of data needs to be saved. Generally, grayscale is divided into the following four methods: component method, maximum value method, average method and weighted average method [6]. Among them, the component method is to use the respective component luminances of R, G, and B in the RGB image as the grayscale value of the grayscale image, which can be used as needed in practical applications. The formula for the component method is as follows:

$$\begin{cases} A_1(i,j) = R(i,j) \\ A_2(i,j) = G(i,j) \\ A_3(i,j) = B(i,j) \end{cases} \tag{1}$$

In the formula, $A_I(i,j)$ and $I = 1, 2, 3$ represent the gray value of each component at the coordinate point (i,j).

3.2 Image Denoising

In aerial images, often due to the influence of dust and small objects on the ground, as well as the camera itself, some irregular patterns such as points and lines are generated. These points and lines are called image noise and cause interference to image processing. Therefore, it is necessary to denoise the image in the preprocessing stage.

On this basis, using the principle of signal analysis, the plane image in the spatial domain is transformed into a frequency image. Among them, the high-frequency

band represents a rapid change in the grayscale of the original image, while the low-frequency band represents a slow change in the grayscale of the original image. In this way, frequency domain filters can be used to remove small interference points that change quickly, thereby achieving the goal of denoising the image. Due to the one-to-one correspondence between images in the time and spatial domains, this operation can also be transformed into convolutional filtering on the image matrix. This article proposes a new method based on convolutional filtering, which adds different weights to the image and adds different weights to the image to form a new image.

At present, the common filtering methods mainly include median filtering, Gaussian filtering, mean filtering and so on. For UAV inspection images, this section mainly uses Gaussian filtering and median filtering. Gaussian filtering is to weight the gray value of each pixel in the image, and then perform a normalization operation, which is mainly used to remove Gaussian noise in the image. Its template function is shown in formula (2).

$$E(i,j) = \frac{e^{\frac{i^2+j^2}{2s^2}}}{2\pi s} \tag{2}$$

In the formula, $E(i, j)$ represents the Gaussian filter function, s represents the standard deviation.

Theoretically, a Gaussian distribution has all domains of non-negative values, which requires an infinite convolution kernel. In fact, only the part within 3 standard deviations needs to be kept, and the other parts can be deleted directly.

The Gaussian function has five important properties, which enable it to play a greater role in different processes. The above properties indicate that Gaussian smooth filtering is a very effective image processing method in the spatial and frequency domains, and it is also a good image processing method [7]. For Gaussian functions, the following four aspects are of great significance:

(1) Gaussian filters have a strong rotational symmetry, therefore, ideal Gaussian filters have consistent appearances in all cases.
(2) We studied the single numerical properties of a Gaussian filter. That is to say, when each signal is subjected to Gaussian filtering, the corresponding filtering weights are different.
(3) The Gaussian function has strong anti-interference ability against time-frequency changes.
(4) By using Gaussian filters, filters of any size can be constructed; This is because there are no restrictions on the left and right sides of the Gaussian function, and this unrestricted situation allows users to expand it.

The smoothness of an image after Gaussian filtering is determined by the standard deviation. Its output is the weighted average of pixels, with pixels closer to the center being weighted more heavily. This algorithm has better smoothness and edge preservation than the average filtering algorithm. Gaussian filter is essentially a low-pass filter, which is a smoothing filter. Follow these steps to perform Gaussian filtering.

Step 1: Move the central element of the associated core directly above the image to be processed.

Step 2: Take the pixel values of the image as the weights and multiply them with the correlation kernel.
Step 3: Accumulate the results obtained from the previous steps to obtain the output.

3.3 Image Enhancement

The purpose of image enhancement is to enhance the information you need and reduce unnecessary information as much as possible, so that the differences between the objects in the image can be enlarged. The enhancement in this chapter is mainly to highlight the details in the image by adjusting the range of pixel values. Many backgrounds in the inspection images are useless information, such as houses, sky, trees, etc. Moreover, due to the problem of light, the object to be detected will appear too dim, which greatly increases the difficulty of identification. So, this section adopts histogram equalization to expand the difference between the detected target and the background and improve contrast [8].

For those cases where the gray values of the image are concentrated in a narrow area, find a way to spread them evenly, and distribute the gray values evenly in each pixel interval by introducing a histogram. In this way, the whole image looks sharper, the contrast is higher, and some details will be more obvious, which is the role of image enhancement.

In order to make the difference between pixel points and other points more obvious, and also facilitate the detection of points, it is necessary to flatten the points. Histogram equalization is a correction method obtained by using the cumulative distribution function transformation method to process images. Histogram equalization refers to maximizing the information quality (entropy) of an image by making the grayscale distribution of its various parts relatively uniform. Its expression is:

$$B = C(e) = (D-1)\int_0^e f(v)\mathrm{d}v, 0 \le e \le D-1 \tag{3}$$

In the formula, B represents the grayscale value after histogram equalization, C represents an increasing and monotonic grayscale transformation function, e represents the original grayscale value of the pixel before enhancement, f represents a probability density function, v represents the hypothetical variable for integration, D represents the gray level before the equalization operation, d represents the derivation symbol.

Histogram equalization is the process of using a grayscale conversion function to transform the original image's histogram into a uniformly distributed grayscale value, and then correcting the original image through the equalized histogram. This method is based on probability theory and utilizes gray point operations for transformation to enhance the image. The conversion function of an image is determined by the cumulative distribution function of the image. By converting this function, a new image is obtained, and the probability density of the image is consistent. When the histogram of the image is evenly distributed, the amount of information contained in the image is the largest, and the image appears relatively clear. The histogram equalization process is as follows:

Step 1: Count the grayscale histogram of the initial image;

Step 2: Accumulate and integrate the image to obtain a uniform distribution of grayscale values in the new image, thereby obtaining the conversion relationship of grayscale values in the image;

Step 3: The third step is to obtain the image processing results after histogram equalization based on the established grayscale conversion relationship.

4 Overhead Line Pin Image Segmentation

Image is an important medium for human information transmission and the main source of understanding the world. In human life, due to the needs of work and research, only a part of the image is needed. How to segment the target image requires image segmentation technology to solve this problem [9]. Image segmentation refers to dividing an image into several non-overlapping sub-regions, so that the features in the same sub-region have a certain similarity, and the features in different sub-regions show obvious differences. Image segmentation is the most important step in the study of images. Whether the work after image segmentation, such as image classification and image analysis, can meet the expected requirements is largely affected by the quality of image segmentation. Due to the complex image segmentation technology, it is difficult to obtain an algorithm suitable for most pictures, and there is no standard for evaluating the quality of the segmentation algorithm, which brings a lot of inconvenience to the specific application of the image segmentation algorithm. An aerial image containing pins can be semantically divided into two parts: foreground (pins) and background (other objects in the frame). Considering the complexity of pictures in natural scenes, it is difficult to find a good and fast traditional image processing method for semantic segmentation of the entire image. However, through a large number of manual annotations, it is still possible to find some algorithms for image region segmentation, among which the image segmentation based on the watershed algorithm is the most effective method. The specific process is shown in Fig. 2.

The watershed segmentation algorithm has many advantages in image segmentation, so it is used in various fields. The main reason is that after segmentation by the watershed segmentation algorithm, a single pixel segmentation line, that is, a watershed, can be obtained, and the watershed segmentation algorithm is highly sensitive to subtle grayscale changes in the image. Therefore, the edge of the object can be accurately located in the image, and finally the segmented area is closed and connected.

5 Defect Identification of Overhead Line Pins

Traditional power devices and fault identification methods often use manual design. For example, the image was preprocessed for enhancement and denoising. By utilizing features such as Harr, invariant moments, and color space, combined with methods such as SVM and cascaded Adaboost, electrical components such as shock absorbers and insulators and their corresponding faults are identified. In addition to the requirement for a large number of expert techniques, this method is often only used for specific classifications and has low scalability. In recent years, with the rapid development of deep learning, it is proposed to use deep learning algorithms to identify power components and corresponding defects. However, most of these researches focus on the identification of components such as insulators and anti-vibration hammers, and the identification of related defects such as insulator missing and self-explosion. The research on pin-level

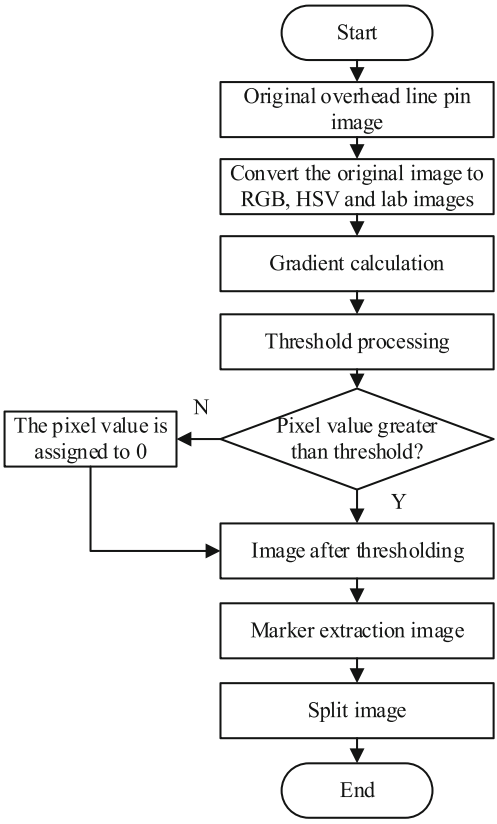


Fig. 2. Image segmentation process of overhead line pins

defects is rare, or even blank [10]. To address these issues, this paper uses the Faster RCNN network to identify pin faults in overhead lines. Faster-RCNN is developed on the basis of Fast-RCNN. The whole model is divided into three parts: the first part uses the hierarchical residual convolution module to extract features, the input is the original image sample, and the output is the feature map of the image. And use a top-down feature map fusion module with horizontal connections and upsampling operations to fuse the deep and shallow feature maps output from the upper part. The second part is the region generation network RPN. The input is the fused high-dimensional feature output from the previous step, and the output is the location score and category score of the region of interest (ROI). The third part sends the ROI position and feature map obtained in the previous steps into the Fast-RCNN network to confirm the category and output the detection result.

5.1 Feature Extraction

The traditional feature extraction network mainly outputs feature maps of multiple scales through the external hierarchical output structure to obtain features of multiple scales,

such as the commonly used feature pyramid structure. The hierarchical residual convolution module directly adds a multi-level structure inside the convolution layer and connects it in the form of residual, which can extract features of multiple scales at a fine-grained level. At the same time, the receptive field of the output features of each layer is increased, so that each position contains richer contextual information. The hierarchical residual convolution module divides the input features into M layers according to channels, denoted as $P_i, i = 1, 2, ..., M$, the resolution of each layer is the same as the input resolution, and the number of channels is the input $1/M$. First, pass the second layer through a 3×3 convolution to get the output of this layer. Afterwards, starting from the third layer, add the features of each layer to the output of the previous layer, and then use 3×3 convolutions are used to obtain the output features of each layer, and the output of the first layer is an identity map. Finally, the output features of each layer are spliced together by channel, and fused through a 1×1 convolution to obtain the final output result. The output $Q_i, i = 1, 2, ..., M$ of each layer is expressed as:

$$Q_i = \begin{cases} P_i, i = 1 \\ G(P_i), i = 2 \\ Q_{i-1} + G(P_i), 2 < i < M \end{cases} \tag{4}$$

where, G represents the transfer function.

Each time the feature undergoes a 3×3 convolution, the receptive field will increase, and the output features of each layer can receive the feature information of each previous layer. This makes the output of each layer contain a combination of different receptive field sizes, so that the fused output features contain feature information of multiple scales and richer contextual information. It can effectively improve the model's ability to recognize small pin targets and the ability to distinguish pins in different parts, and finally improve the detection effect of the model. The more layers the module is divided into, the larger the output receptive field and the richer the acquired features. However, the complexity of the model will increase accordingly. Here, the number of divided layers is finally set to 4 to measure the accuracy and model complexity of the model.

5.2 Delineation of the Region of Interest

The traditional target detection method generally uses a sliding window to traverse the image to generate a candidate frame, and then use some classifiers to complete the classification of the candidate frame. This method has poor real-time performance and poor effect. The RPN network is used in Faster-RCNN to generate candidate boxes. First, RPN generates anchors (Anchors) for each position of the feature map. Anchors consist of several fixed scales and several fixed aspect ratio rectangular boxes. Second, these Anchors are input into the RPN classifier, and the classifier determines whether the Anchors contain targets. If they contain targets, the Anchors are retained, otherwise the Anchors are eliminated. And perform bounding box regression on Anchors to output regions of interest (ROIs).

5.3 Classification of Pin Defects

The low accuracy of the "single-order" detector is attributed to the class imbalance of the candidate regions. That is, the target to be recognized in the picture usually only occupies a small part of the whole picture, so that most of the candidate windows generated in the early stage belong to the negative class (background), and only a very small part contains the foreground object. The dominance of the background class obscures the effect of the candidate frame containing the recognized object, so that the training process cannot fully learn the required information. For Fast-RCNN, the Region Proposal Network (RPN) it contains performs simple binary classification of candidate regions in advance, so that the candidate windows belonging to the background will be greatly reduced. This reduces the impact of class imbalance of candidate regions to some extent, but the complexity of its operation slows down the recognition speed.

Taking binary classification as an example, Anchors generated by each RPN are assigned a positive or negative label. The cross entropy loss (Focal loss) of the i Anchors is shown in formula (5).

$$\begin{cases} R(\alpha_i, \beta_i) = -\log \alpha_i, \beta_i = 1 \\ R(\alpha_i, \beta_i) = -\log(1 - \alpha_i), \beta_i = 0 \end{cases} \tag{5}$$

In the formula, $\beta_i \in \{0, 1\}$ is the sample label of the i Anchors, and $\beta_i = 1$ indicates that the Anchors and a bounding box have a larger IoU, which is a positive sample. $\beta_i = 0$ indicates that the anchors are negative samples, α_i is the score of the i Anchors belonging to the positive samples, and $1 - \alpha_i$ is the score of the i Anchors belonging to the negative samples. When Anchors are positive samples, the classification loss decreases with the improvement of sample detection accuracy. When Anchors are negative samples, the classification loss decreases with the decrease of sample detection accuracy.

In order to learn more information from the candidate region containing the target, a weight control coefficient w is added to the original cross-entropy loss function $R(\alpha_i, \beta_i)$, namely:

$$\begin{cases} \hat{R}(\alpha_i, \beta_i) = -w(1 - \alpha_i)^{\lambda} \log \alpha_i, \beta_i = 1 \\ \hat{R}(\alpha_i, \beta_i) = -(1 - w)\left[-(1 - \alpha_i)^{\lambda}\right] \log(1 - \alpha_i), \beta_i = 0 \end{cases} \tag{6}$$

In the formula, w is the balance factor, which is used to balance the contribution of positive and negative samples to the classification loss. The smaller w is, the smaller the loss weight of negative samples is. λ is an adjustment factor, which is used to adjust the loss weight of simple negative samples and difficult negative samples. When $\lambda = 0$, Focal loss is the cross entropy loss. The larger λ is, the greater the loss of difficult negative samples. For the i Anchors, when the Anchors are positive samples, the larger α_i is, the smaller $\hat{R}(\alpha_i, \beta_i)$ is, that is, the loss of the i positive sample decreases with the increase of α_i. When the Anchors are negative samples, the larger α_i is, the larger $\hat{R}(\alpha_i, \beta_i)$ is, that is, the loss of the i difficult negative sample increases with the increase of α_i. Therefore, by weighting parameters w and λ, Focal loss can effectively improve the loss weight of difficult negative samples.

Secondly, the selection criteria of Anchors by RPN are improved, and the number of all Anchors in an image is reserved and input into RPN as a training batch. In Fast-RCNN, the classification loss of the improved RPN is shown in formula (7):

$$Z = \frac{\sum_{i=1} \hat{R}(\alpha_i, \beta_i)}{N} \tag{7}$$

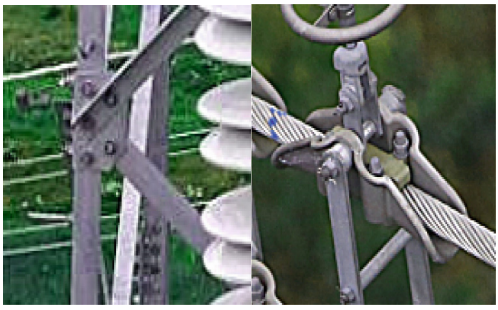
In the formula, N is the number of all positive and negative sample Anchors of an image in each batch input to the RPN, β_i is the i Anchor and the corresponding positive and negative samples; $\hat{R}(\alpha_i, \beta_i)$ is the Focal loss of the i Anchors.

6 Method Application Testing

6.1 Experimental Dataset

To verify the effectiveness of the method for identifying pin defects in overhead transmission lines based on drone multi-sensor data, this project plans to use the original dataset obtained during drone detection of transmission lines, and some of the images of overhead line pins are shown in Fig. 3.

(a) Pin defect 1 (b) Pin defect 2



(c) Pin defect 3 (d) Pin defect 4

Fig. 3. Schematic diagram of overhead line pins

The original image contains five types of targets, such as no defect, missing pins, pin prolapse, missing spacers, and pin deformation. They are unevenly distributed in the

image, and there is a situation where one image contains multiple targets and multiple targets. In the original data set of this paper, the number of samples with no defects, missing pins, pin prolapse, and missing gaskets is relatively large, while the number of samples with pin deformation is relatively small. Therefore, this paper generates artificial samples according to the number of samples of different types of targets, and finally makes the number of samples of each type equal. The specific number of samples is shown in Table 1.

Table 1. Experimental dataset distribution

Type	Sample type:	Number of samples
No defect	Training sample	235
	Test sample	241
Lack of pin	Training sample	362
	Test sample	354
Pin out	Training sample	287
	Test sample	241
Lack of gasket	Training sample	300
	Test sample	292
Pin deformation	Training sample	252
	Test sample	238

6.2 Image Feature Extraction Results

Taking the partial empty line pin image as an example, the image features extracted by the hierarchical residual convolution module are shown in Table 2.

6.3 Delineation of the Region of Interest

The RPN network is used to delineate the region of interest (target) in the overhead line pin image, and the results are shown in Fig. 4.

6.4 Analysis of Detection Method Performance

Use training samples to train the pin defect detection method based on Faster-RCNN (the research method) and the online fault detection method based on DKPCA algorithm for the fixed wing UAV with multiple operating conditions (the method of reference [2]), then use the trained method for defect detection, and finally compare it with the actual results to calculate the AP value. The AP value is composed of precision ψ and recall ζ.

Accuracy ψ refers to the proportion of truly defective samples identified as defective samples. The recall rate ζ can also be called the detection rate, which refers to the

Table 2. Image feature extraction results

Image	Characteristic scale 1	Characteristic scale 2	Characteristic scale 3	Characteristic scale 4
1	0.2456	2.5852	2.8558	6.8652
2	0.3557	2.9878	1.4422	4.4233
3	2.5452	3.4552	0.5532	0.0123
4	0.0335	0.8465	2.5474	1.2222
5	0.6852	0.7425	0.0532	3.5452
6	1.5233	3.8641	0.0014	6.4502
7	0.1523	1.1255	2.4522	4.3212
8	2.1441	1.5221	3.2152	0.0212
9	0.1521	0.3845	3.2001	0.4822
10	0.0662	0.7452	3.7220	0.3435
11	1.5922	1.1257	4.1254	1.2122
12	0.1256	2.6522	2.0531	2.2020
13	0.0865	0.0865	0.1501	2.2127
14	0.1245	3.4571	0.5320	0.0422
15	1.8745	2.5455	0.5643	0.7429
16	2.5653	0.0162	0.2933	4.1233
17	0.0652	0.0252	2.9872	2.6513
18	0.5381	0.0142	3.1225	2.8622
19	1.6239	3.5422	4.2152	3.4221
20	2.2250	0.4522	1.6422	0.5274

proportion of the identified true defective samples to all defective samples. The specific calculation formula is:

$$\psi = \frac{x}{x+y} \tag{8}$$

$$\zeta = \frac{x}{x+z} \tag{9}$$

In the formula, x represents the number of defects detected correctly, y represents the number of defects detected by mistake, and z represents the number of defects that were missed.

Recall and precision vary for different confidence thresholds. Taking ζ and ψ as the horizontal and vertical coordinates, respectively, an ψ-ζ curve can be obtained. The AP value is the area of the area enclosed by the curve and the horizontal and vertical axes. The value is between 0 and 1. The larger the value, the higher the accuracy of the detection method. The AP value can comprehensively and comprehensively evaluate the performance of the algorithm.

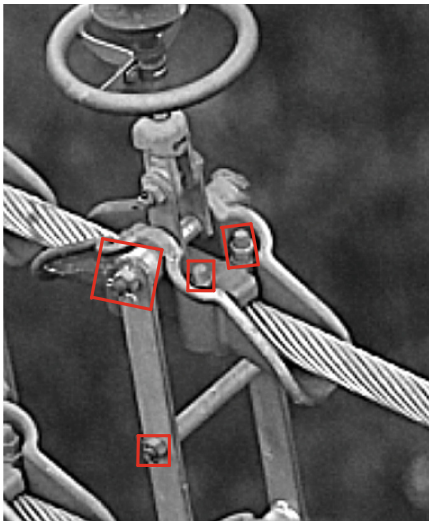


Fig. 4. Example of delineating the region of interest

The AP values of the two detection methods are shown in Table 3.

Table 3. AP value of two detection methods

Type	AP value of the research method	AP value of the method of reference [2]
No defect	0.9555	0.7525
Lack of pin	0.9254	0.8754
Pin out	0.9712	0.7012
Lack of gasket	0.9314	0.8384
Pin deformation	0.9252	0.8292

It can be seen from Table 3 that under the application of the research method, the AP values of the 'five types of samples are all above 0.9, indicating that the research method has high accuracy.

7 Conclusion

In addition to working outdoors in harsh environments, most power fittings in the power system also need to withstand the external mechanical load tension and the power load inside the power system for a long time. This makes the pins on the fittings prone to defects such as missing or loose, affecting the stable operation of the power grid. To this end, a method for detecting pin defects in overhead lines based on multi-sensor data

acquisition by UAV is proposed. The pin image is collected by UAV multi-sensor, and the image is preprocessed. The watershed algorithm is used to segment the background part of the overhead line pin image, and the Faster-RCNN network is used to complete the identification of the overhead line pin defect. The method is finally tested by application, and the obtained AP values are all above 0.9, which proves the effectiveness of the method.

References

1. Tang, F., Gao, Q., Du, Z.: Algorithm of object localization applied on high-voltage power transmission lines based on line stereo matching. Opt. Eng. **60**(2), 023101 (2021)
2. Liang, S., Zhang, S., Zheng, X., et al.: Online fault detection of fixed-wing UAV based on DKPCA algorithm with multiple operation conditions considered. Xibei Gongye Daxue Xuebao/Journal of Northwestern Polytechnical University **38**(3), 619–626 (2020)
3. Le, V., Yao, X., Hung, T.B., et al.: Series DC arc fault detection based on ensemble machine learning. IEEE Trans. Power Electron. **35**(8), 7826–7839 (2020)
4. Li, H., Li, G., Ye, Y., et al.: A high-efficiency acquisition method of LED-multispectral images based on frequency-division modulation and RGB camera. Optics Communications **480**(4), 126492 (2021)
5. Swamy, A., Srinath, S.: POS Tagging and NER System for Kannada Using Conditional Random Fields. International J. Information Retrieval Research (IJIRR) **11**(4), 1–13 (2021)
6. Yang, Y., Zhang, P.: A novel bond wire fault detection method for IGBT modules based on turn-on gate voltage overshoot. IEEE Trans. Power Electron. **36**(7), 7501–7512 (2021)
7. Zheng, H., Wang, R., Yang, Y., et al.: Intelligent fault identification based on multisource domain generalization towards actual diagnosis scenario. IEEE Trans. Industr. Electron.Industr. Electron. **67**(2), 1293–1304 (2020)
8. Wang, L., Chang, X., Ren, W.: Color image enhancement simulation based on weighted histogram equalization. Computer Simulation **38**(12), 126–131 (2021)
9. Yanxia, C., Yanyan, X., Tierui, Z., et al.: Threshold image target segmentation technology based on intelligent algorithms. Компьютерная оптика **44**(1), 137–141 (2020)
10. Jiang, C., Fang, Y., Zhao, P., et al.: Intelligent UAV identity authentication and safety supervision based on behavior modeling and prediction. IEEE Trans. Industr. Inf.Industr. Inf. **16**(10), 6652–6662 (2020)