

PAPER • OPEN ACCESS

Pin Defect Detection Method of UAV Patrol Overhead Line Based on Cascaded Convolution Network

To cite this article: Gu Chaoyue *et al* 2020 *J. Phys.: Conf. Ser.* **1659** 012021

View the [article online](#) for updates and enhancements.



ECS The Electrochemical Society
Advancing solid state & electrochemical science & technology

239th ECS Meeting with IMCS18

DIGITAL MEETING • May 30-June 3, 2021

Live events daily • Free to register

18th

Register now!

Pin Defect Detection Method of UAV Patrol Overhead Line Based on Cascaded Convolution Network

GU Chaoyue^{1,a}, LI Zhe^{1,b}, SHI Jintao^{1,c}, SHENG Gehao^{1,d} and JIANG Xiuchen^{1,e}

¹Department of Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China;

^aandrew_gcy@163.com; ^bzhe_li@sjtu.edu.cn; ^csjtu_sjt@sjtu.edu.cn;

^dshenghe@sjtu.edu.cn; ^excjiang@sjtu.edu.cn

Abstract. Pin plays the role of fixing power equipment on the overhead line. Once it is missing, it will lay a serious hidden danger for the normal operation of the overhead line. In order to improve the efficiency of UAV patrol transmission line and improve the detection rate of pin defect of transmission line, this paper proposes a pin defect detection method based on cascaded convolution network. In view of the complex background of inspection image and the small size of pin, the overall detection method is divided into two parts: positioning and diagnosis. Firstly, all fastener positions including pins are located by the improved Faster-RCNN network, and then RetinaNet network is cascaded to diagnose the defects of the fastener. The difficulty of learning is decreased and the generalization ability is improved in this way. Finally the experiment shows that this method can effectively detect the pin defects in the UAV patrol image.

1. Introduction

Due to its high efficiency, flexible operation and easy to carry, UAV patrol inspection is gradually popularized and applied in various power grids across the country. By operating UAV for remote inspection, inspectors take transmission line images at fixed positions, and then analyze the images to further complete troubleshooting and timely maintenance of transmission line faults, effectively improve inspection efficiency, reduce grid operation costs, and the massive data generated by inspection is conducive to recording and storage, which is conducive to the construction of power big data.

Owing to the large scale, complex background and large number of aerial inspection images, the manual inspection workload is large, the efficiency is low, and the rate of missed inspection is high [1]. With the development of machine learning and deep learning technology, by the help of appropriate and reliable algorithms, the defects in inspection images can be identified and located automatically, which can further improve the automation of UAV inspection and greatly improve the efficiency of inspection. In recent years, researchers have done many useful researches[2-6] on the defect detection of UAV inspection image of transmission line. In reference [2], firstly, the candidate window of insulator is extracted by using the binary normed gradient classifier, then the window set including insulator target is obtained by using the convolution neural network accurate recognition, finally, the window set with high overlap is combined by weighted iteration to locate insulator. In reference[3], yolov2 network was trained to learn the insulator features in the inspection image, and the insulator defects were diagnosed by combining edge detection, line detection and other image



processing methods. reference [4] combined with histogram equalization, morphological processing and RGB color model, through the comparison of normal and rusty conditions of shock hammer practice on the detection of rust defects. In reference[5], the Faster RCNN algorithm is used to detect the small power components in the UAV patrol image, and it is compared with DPM and sppnet algorithm respectively. In reference[6], firstly sharpen the inspection image, then use the yolov3 network to intercept the anti bird thorn part of the inspection image, and finally use the resnet152 network to complete the fault detection of the anti bird thorn area.

There are many important parts on the transmission line, such as insulator, grading ring, shock hammer and so on. Generally, these parts are fixed by pins and fittings to prevent them from maloperation and affect the reliable operation of the transmission line. As the transmission line is exposed all year round, it will inevitably suffer from severe environment such as rainstorm, high temperature, lightning, bird damage, etc., which may cause pin to come out or fall off, seriously affecting the normal operation of other parts[7].

Due to the large scale of inspection image, complex background and small pin size, the existing target detection algorithm is difficult to effectively detect and accurately locate the defective pin. In this paper, a pin defect detection method based on cascaded convolution network is proposed, which divides the small target detection task into two steps: location and detection. In the first stage, the improved Faster RCNN[8] network is used to locate the fastener position containing the pin. In the second stage, the RetinaNet[9] network is used to determine whether the pin part has defects. In the two stages, the network uses the Different, made targeted improvements, adjusted some parameters and network structure. Experiments show that this method has better detection effect and more accurate positioning ability for pin defects in aerial line inspection image of UAV.

2. General overview

The overall flow of this detection method is shown in Figure 1. Firstly, the all fastener locations containing pins are located by the improved Faster-RCNN network in one phase, then automatically clipped and resized, and sent to the RetinNet network in the second phase to determine whether the pins are defective. Finally, a rectangular frame is drawn in the inspection image to locate the defective pins.



Figure 1. General flow chart of detection method

2.1. Positioning fastener

The input of the first phase of the network is a high-pixel UAV patrol image, and the task is to locate all fastener locations containing pins in the image. Because the input image has large scale, high resolution, complex background, small scale of fastener, difficult to identify, and high accuracy requirements for the algorithm, Faster-RCNN algorithm is selected to train and detect at a larger image scale, and feature Pyramid Network[10] (Feature Pyramid Networks, FPN) improves the detection of small targets, and adjusts the initial anchor frame size and shape after clustering using K-means[11] algorithm to improve the overall performance of the algorithm.

2.2. Defect Discrimination

Based on the fastener location of the first phase network output, the fastener is automatically clipped and resized from the original image, and is used as the input of the second phase network. The task of the second phase of the network is to determine whether the pins in the fastener fall off or fall out and locate the defective pin position. Considering that first-stage network may have multiple outputs,

second-stage network will be called repeatedly, which requires a faster detection. RetinaNet algorithm, which has relatively simple structure, fast recognition rate and high accuracy, is selected.

3. Detection algorithm

3.1. Location network based on Faster-RCNN

Faster RCNN network is a two-stage target detection network with high accuracy. It is composed of pre-feature extraction network CNN, area recommendation network RPN and detection network RCNN. CNN extracts the feature map of image. RPN output the category score and suggesting bounding box that may contain the detected object through the feature map. Finally, RCNN output the target classification confidence and positioning rectangle position based the suggesting bounding box and the feature map.

In order to improve the detection effect of small targets, we build a feature pyramid structure, which integrates the multi-scale convolution features from the bottom to the top and from the shallow to the deep. On the basis of retaining the deep abstract semantics, shallow features are learned. The front network uses the deep residual network ResNet101[12] to replace VGG16 network. The problem of gradient disappearance or explosion and network degradation of traditional networks when the model is too deep is overcome by using its short-circuit connection. The K-means clustering algorithm is used to calculate the target scale distribution of data set and optimize the initial anchor

3.2. Discrimination network based on Retinanet

Retinanet, as a single-stage detection network, has a relatively simple and clear structure, as shown in Figure 2. Feature extraction network CNN extracts convolution features of high-resolution input image and outputs feature images of different scales. The feature pyramid network FPN uses the feature graphs {C3, C4, C5} of each layer of CNN network, and then performs a series of simple operations such as up sampling, 1×1 convolution, weight addition (see the FPN part in Figure 2 for details), and establishes the feature pyramid structure {P3, P4, P5, p6, P7}, the purpose is to fuse the features of each scale of the feature map. On the basis of retaining the deep semantic, more shallow features are extracted, which is helpful for the recognition of small targets. The target detection network is composed of two full convolution networks, which are used to calculate the confidence of target classification and to regress the location of rectangular box.

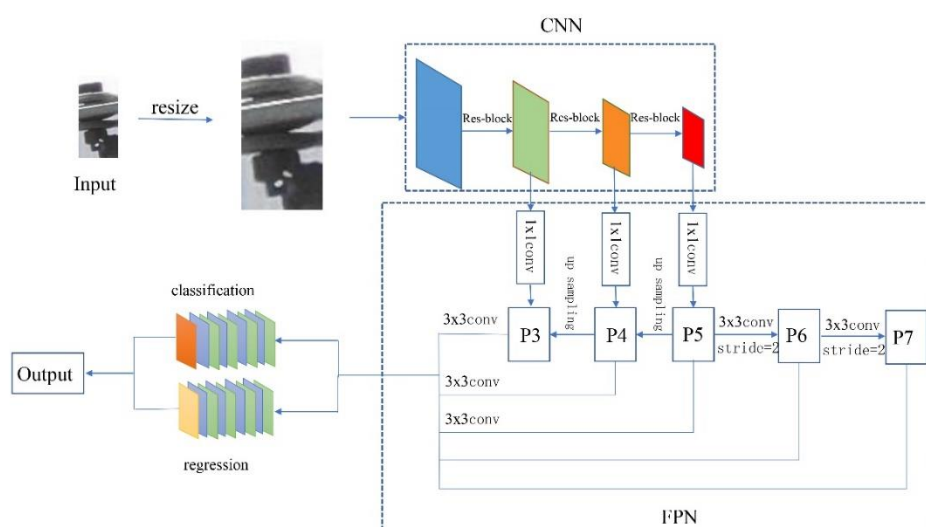


Figure 2. Network structure of RetinaNet

One of the important reasons why the single-stage detection algorithm has weak detection ability is the problem of category imbalance. Especially for small targets, thousands of candidate boxes only contain a very small number of positive samples, and the proportion of negative samples is too high, so that it dominates the training direction, the model has not been effectively trained, and the detection effect is poor. To solve this problem, focal loss is used in RetinaNet, and the specific formula is shown in formula (1), (2), where γ is the dynamic scaling factor, in this paper, it is 2, α is the balance variable, in this paper, it is 0.25, $y \in \{0, 1\}$ is taken to represent the category label, $P \in [0, 1]$, to represent the probability that the output category of the model is positive.

For easy to classify samples or negative samples, the p_t value is large, so the weight is small. For hard to classify samples or positive samples, the p_t value is small, so the weight is large, which makes the model pay more attention to such samples in training. Therefore, focal loss alleviates the imbalance of positive and negative samples to a certain extent, and also solves the imbalance of difficult and easy samples.

$$FL(p_t) = -\alpha_t (1-p_t)^\gamma \log(p_t) \quad (1)$$

$$p_t = \begin{cases} p & y=1 \\ 1-p & y=0 \end{cases} \quad \alpha_t = \begin{cases} \alpha & y=1 \\ 1-\alpha & y=0 \end{cases} \quad (2)$$

4. Experiment

4.1 data set establishment and evaluation index

Since there is no relevant public data set available at present, the sample used in this experiment is the high-resolution image of UAV inspection overhead line provided by some provincial power companies. In order to improve the generalization ability of the model, enhance the detection performance of patrol inspection images under different shooting angles, lighting conditions and shooting articulation, we have made data enhancement for the training samples, including horizontal inversion, Gaussian blur, overall brightness adjustment, and expanded the training set to twice. The specific data set composition is shown in Table 1. In order to facilitate subsequent training and testing, all data are made into Pascal VOC format.

Table 1. Data set of Pin

	Training Pic	Target number	Testing Pic	Target number	Image size
First Stage	3000	13694	248	1073	4000*3000
Second Stage	1909	1909	210	211	600*400

The evaluation index of this experiment is AP value. The definition of recall rate is the ratio of the defect samples detected by the algorithm to the defect samples in all test samples, which indicates the detection ability of the algorithm. The definition of accuracy is the ratio of the actual real defect samples detected by the algorithm, which indicates the accuracy of the algorithm. The recall rate and accuracy generally increase with the lower confidence level. The overall relationship is inverse. The AP value is the area enclosed by the curve and the coordinate axis with recall and precision as the horizontal and vertical coordinates respectively. The closer the area is to 1, the better the comprehensive performance of the algorithm. AP value can better reflect the comprehensive detection performance and generalization ability of the model

4.2 Model training and testing

The hardware configuration of this experiment is as follows: CPU: Intel Core i7-8700k, 3.70ghz main frequency. GPU: NVIDIA gtx1080ti, 11g memory. The software platform is Linux Ubuntu 18.04.

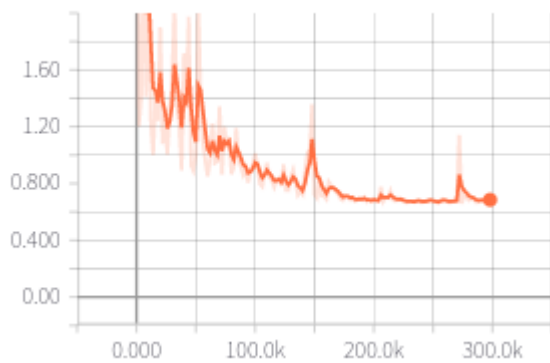
In the experiment, we use Tensorflow framework to build the cascaded convolution network model, and the two-stage pre feature extraction network adopts the fine-tune [13] strategy, which can improve

the training convergence speed to a certain extent by fine tuning the network parameters pre trained by the Imagenet dataset. The optimization strategies are the stochastic gradient descent with momentum[14] (SGDM), where momentum is set to 0.9. The detection threshold is set to 0.7, dropout, batch processing parameters, non maximum suppression parameters and other parameters are consistent with the original network.

Because of the different tasks between the two-stage network, there is a large difference between the network parameter setting and the training strategy. For the first stage network, the input is a high-resolution inspection image, and the output is a rectangular frame of all fastener positioning including pins. The pre convolution neural network model uses the depth residual network resnet101, uses large-scale training, resizes the short side of the image to 900px, and the long side to 1500px, and increases the scale of the feature image. After testing, the initial learning rate is 10^{-3} , after training 100000 steps, it decays to 10^{-4} , after training 180000 steps, it decays to 10^{-5} . The convergence trend of loss is as shown in Figure 3 (a). The overall trend shows a downward trend to 200000 steps, and then it basically converges, and enters a stable state, with good effect.

The training sample of the two-stage network is the image of fastener part captured from the high-resolution image of UAV inspection overhead line, and the short side is resized to 400px and the long side to 600px. Resnet50, a relatively shallow residual network, is used in the pre neural network model, with an initial learning rate of 10^{-3} . After 150000 steps of training, it decays to 10^{-4} . After training 200000 steps, it decays to 10^{-5} . In the training process, the total loss of the network decreases as shown in Figure 3 (b). After the number of training steps reaches 250000, the network basically converges to a stable state with good effect.

LOSS/total_loss



(a) Downward trend of first stage network's Loss

LOSS/total_loss



(b) Downward trend of second stage network's Loss

Figure 3. Loss trend chart

Due to the small size of the defective pin, the IOU value will be greatly reduced if the position of the predicted positioning frame is slightly deviated. After several tests, the threshold value of IOU is set to 0.3. When the IOU value of the two frames is above this threshold value, it can be more accurately located to the specific location of a defective pin on the tower, so that the staff can easily carry out troubleshooting and maintenance, basically meet the actual application needs, so it can be considered that the detection is successful.

4.3 analysis and comparison of results

4.3.1 Fastener positioning. Import the ckpt model file generated by the training, input the test set for testing, part of the test results are shown in Figure 4, and accurately locate the fasteners in the test sample.



Figure 4. Detection results of first stage network

The first stage network algorithm of this method is based on Faster RCNN network, but has made some improvements. In order to verify the feasibility of the improvement strategy, different strategies are used to train and test the network. The specific detection results are shown in Table 2. Strategy 1 is the initial fast RCNN network. In strategy 2, resnet101 is used as the pre feature extraction network instead of vgg16, the AP value is increased by 1.1%, and the detection rate is basically unchanged. Strategy 3 builds feature pyramid and integrates multi-scale features. Compared with strategy 2, AP value increased by 3.1% and detection time increased by 32.2ms/piece. Compared with strategy 3, strategy 4 improves AP value by 3.3% and detection time by 52ms / piece. This is because it expands the scale of training and testing, and increases the size of feature map. Although it increases some computing power, it effectively improves the detection rate of small targets. In strategy 5, K-means clustering algorithm is used to optimize the initial anchor frame, AP value is increased by 1.3%, and detection rate is slightly improved. The above experiments show that the improved Faster RCNN strategy proposed in this paper can effectively improve the comprehensive detection performance of the target in this paper.

Table 2. Algorithm detection effect under different strategies

Strategy	Backbone	FPN	Large training scale	Optimize initial anchor	AP
1	VGG16				0.801
2	ResNet101				0.812
3	ResNet101	✓			0.843
4	ResNet101	✓	✓		0.876
5	ResNet101	✓	✓	✓	0.889

4.3.2 Pin defect identification. Import the CKPT model file generated by the training, input the test set for testing, and part of the test results are shown in Figure 5(the mark in the figure about pin_defect_0x01 means falling off, pin_defect_0x02 means prolapses), and diagnose whether the pin in the fastener falls off or prolapses.

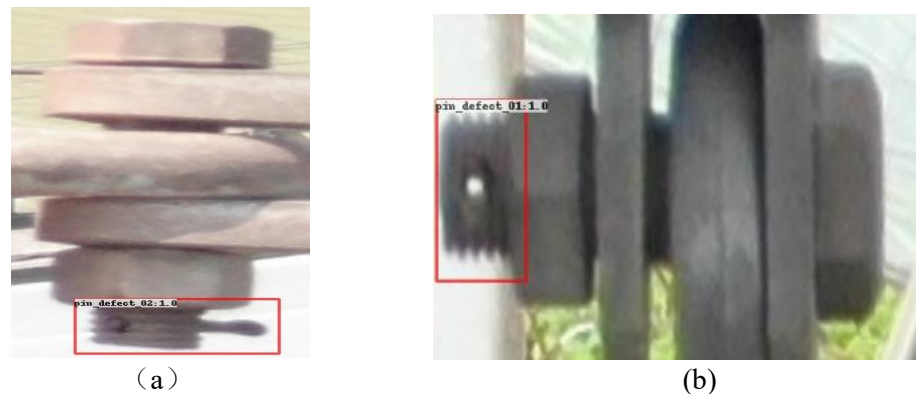


Figure 5. Detection results of second stage network

Because the two-stage network will be called repeatedly, it requires higher speed. The increase of training and detection scale will reduce the detection rate, while the small training and detection scale may lead to the network under fitting and reduce the detection accuracy. Through experiments, the influence of scale size on network accuracy and speed is shown in Table 3. It can be seen from table 3 that the detection rate decreases with the increase of image scale. The influence of scale on the detection accuracy is generally inverted U-shaped, because if the scale of convolution layer after pooling is too small, the network fitting ability will decline, affecting the accuracy; if resize to large size, the quality of image will be too low, and the network learning efficiency will be reduced. Therefore, when the training and detection scale is 400*600, the detection accuracy is high, the detection speed is fast, and the comprehensive detection effect is the best.

Table 3. The influence of different scales on the comprehensive performance of network

Training and testing scale	mAP	Detection time/ms
450*300	0.812	84.0
600*400	0.861	98.0
750*500	0.785	151.5
900*600	0.792	188.7

5. Conclusion

In this paper, a method for pin defect identification in the inspection of transmission line image is proposed. By means of cascaded network, the improved Faster-RCNN network is used to locate all fasteners containing pins, and then the pin state is diagnosed and the position of the pin defect is further located through RetinaNet network. In the test data set, the AP value of the first stage location network reach 0.889, and the AP value of the second stage location network achieve 0.861, which indicates that the network has better detection performance. On the other hand, the flexibility of each stage of the network is strong, and the appropriate strategy or parameters can be selected according to the actual requirements.

References

- [1] Li Junfeng, Wang Qinruo, Li min. Image Recognition of Power Equipment Based on Deep Learning and Random Forest [J]. High Voltage Technology, 2017, 43(11): 3705-3711.
- [2] Peng Xiangyang, Liu Yang, Wang Ke, et al. Using Convolution Neural Network to Locate Insulator Automatically[J]. Journal of Wuhan University (Information Science Edition), 2019, 44(04): 563-569.

- [3] Lai Qinpin, Yang Jun, Tan Bendong, et al. Insulator Automatic Identification and Defect Diagnosis Model Based on YOLOv2 Network [J]. China Power, 2019, 52(07): 31-39.
- [4] Song Wei, Zuo Dan, Deng Bangfei, et al. Detection of Corrosion Defects of Shockproof Hammer for High Voltage Transmission Line [J]. Journal of Instrumentation, 2016, 37(S1): 113-117.
- [5] Wang Wanguo, Tian Bing, Liu Yue, et al. Research on Recognition of Small Power Components in UAV Inspection Image Based on RCNN [J]. Journal of Earth Information Science, 2017, 19(02): 256-263.
- [6] Lin Zhicheng, Miu Xiren, Jiang Hao, et al. Identification and Fault Detection of Bird Stab Proof Components of Transmission Lines Based on Deep Convolution Neural Network [J / OL]. Power System Technology: 1-11[2020-02-20]. <https://doi.org/10.13335/j.1000-3673>. pst.2019. 1775.
- [7] Lv Zhining. Summary of Common Fault Analysis and Detection Methods for Transmission Lines [J]. Automation and Instrumentation, 2020(01): 161-164+168.
- [8] Ren S, He K, Girshick R, et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2015, 39(6): 1137-1149.
- [9] Lin T Y, Goyal P, Girshick R, et al. Focal Loss for Dense Object Detection[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, PP (99):2999-3007.
- [10] T. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan and S. Belongie. Feature Pyramid Networks for Object Detection[C]// 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, America: IEEE, 2017: 936-944.
- [11] Hartigan J A, Wong M A. Algorithm AS 136: A K-Means Clustering Algorithm[J]. Journal of the Royal Statistical Society, 1979, 28(1): 100-108.
- [12] K. He, X. Zhang, S. Ren and J. Sun. Deep Residual Learning for Image Recognition[C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, America: IEEE, 2016: 770-778.
- [13] S. J. Pan and Q. Yang, "A Survey on Transfer Learning," in IEEE Transactions on Knowledge and Data Engineering, vol. 22, no. 10, pp. 1345-1359, Oct. 2010.
- [14] Ketkar N. Stochastic Gradient Descent[M]. Berkeley: Apress, 2017.