

Inspect Defect of Power Equipment via Deep Learning Method

Qi Zhang¹, Qingsong Cai^{1*}, Jianhui Zheng²

¹School of Computer, Beijing Technology and Business University

²School of Software, Tsinghua University

qizhang16@yeah.net, caiqs@btbu.edu.cn, zjh19@mails.tsinghua.edu.cn

Abstract- Efficient defect detection is crucial for maintaining a stable power system. One method is to check the status of electrical equipment through images or videos taken by UAVs (unmanned aerial vehicles), overcoming various difficulties of manual inspection in complex environments such as high-altitude, snowy or rainy days, etc. Unfortunately, it is still a difficult challenge to process the data returned by UAVs automatically and effectively. This study proposes a defect detection algorithm to process UAV images of power equipment based on Faster-RCNN. Additionally, we use multiple optimization strategies to improve the performance of the benchmark algorithm, including feature pyramid network, deformable convolution, online hard example mining, and data augmentation. Our results show that the average accuracy of experiments obtains nearly 60% improvement compared to that naive deep learning target detection algorithm. Meanwhile, this method is suitable for various defects of power equipment with high accuracy, thus benefiting thousands of households.

Keywords— power equipment; defect detection; UAV photography; deep learning

1 Introduction

Periodical inspection, troubleshooting, and timely repair are essential to the stability of the power system, especially to the long-distance high-voltage transmission circuit with complex topographical distributions. The common defects of electric devices are shown in Fig.1, including foreign matters in the tower, damage of the insulator, abrasion of wires, and corrosion of hardware. If these common electrical equipment defects are not found and fixed in time, it will cause extensive equipment damage and paralyze the power grid. To guarantee the steady operation of these circuit devices, unmanned aerial vehicle (UAV) photography has become the top choice for an electric system inspection. UAVs can capture images or videos of electric devices in a complex environment. In contrast, manual inspections often face tremendous challenges under which. And then, we can make out the defect type and position by processing these images or videos. Therefore, intelligent analysis of power equipment defects through images or videos obtained by UAVs plays a significant role in improving fault detection and repair efficiency.

Over the years, tremendous efforts have been made to process the images obtained by UAVs. In general, these methods can be roughly

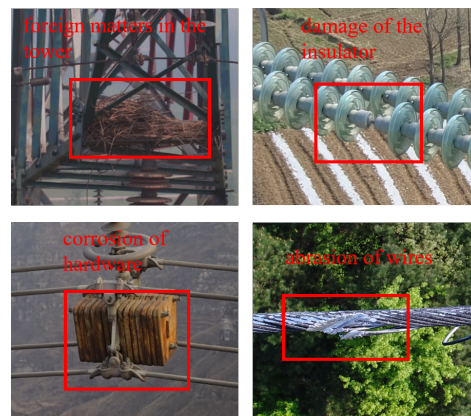


Figure 1: Common defects of power equipment

divided into three categories: The first and the most commonly used one is manual analysis [1], which is to inspect and analyze graphics or video information manually, bearing the highest accuracy. But, it also has obvious drawbacks. The high demand for the quantity and quality of professionals makes it expensive due to the vast amount of power infrastructure relying on manual inspection. Meanwhile, inspecting images manually for an extended period can cause fatigue and reduce judgment accuracy, thereby omitting some equipment defects and causing safety hazards. The second one is the traditional image-based defect detection method that inspects power lines by analyzing infrared images [2]. This method has satisfactory performance in detecting defects that lead to large temperature gradients, such as local high temperatures caused by wire corrosion. Nevertheless, not all power failures will cause temperature changes in power equipment, and infrared image analysis has its limitations. The third one is to use deep learning methods, which utilize data-driven training of convolutional neural networks to locate and classify defects [3]. Deep learning methods can automatically extract image features [4], simplify the image preprocessing process, and effectively improve target detection efficiency compared with traditional image fault detection methods.

However, the existing deep learning methods still have some shortcomings in solving these problems, such as insulator identification methods [5], or power transformer fault diagnosis [6], they can achieve good results under limited data sets. But these methods can only have satisfactory performance for specific problems, which do not work for other kinds of defects. Because Faster-RCNN [7] is widely used in the industry and has high stability, strong versatility, which can relatively accurately identify multiple defects. This paper uses the Faster-RCNN algorithm as the benchmark algorithm. Then we propose an automatic detection method that can recognize

*Corresponding author.

DOI reference number:10.18293/SEKE2021-082

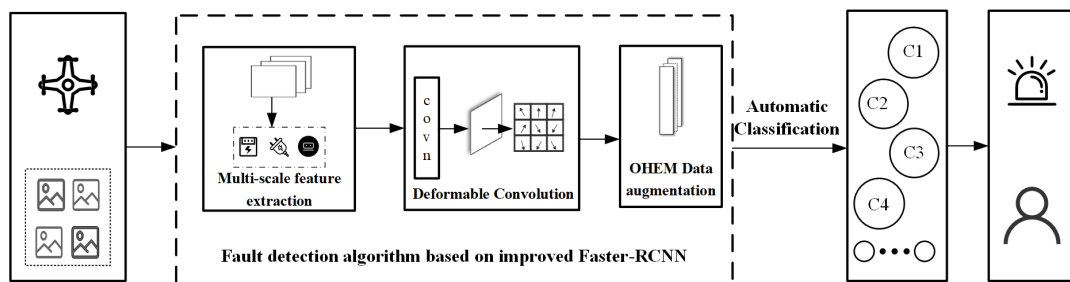


Figure 2: Work flow of the defect detection algorithm based on improved Faster-RCNN

various defects in power equipment.

Although Faster-RCNN is already a relatively mature detection algorithm, defect detection in power systems still faces many challenges: Firstly, there are so many types of defects to be detected that it is difficult for the algorithm to locate them accurately in UAV images. Secondly, complex and diverse backgrounds of pictures will bring annoying noise to the target detection algorithm, thus boosting the robustness of the algorithm in different scenarios must be taken into consideration. Thirdly, UAVs often shoot objects from different heights and angles, so the captured images' target has different sizes and angles. The recognition algorithm needs to be able to detect targets at multiple scales.

This paper proposes a defect detection method for power equipment aerial images based on the improved Faster-RCNN algorithm. The system framework is shown in Fig.2: UAV aerial photography obtains videos or images of power equipment in different environments and transmits these images to the server after relevant processing. The improved Faster-RCNN target detection algorithm can automatically detect defects in power equipment images and automatically classify them, and then professionals will troubleshoot and repair the defect.

According to the characteristics of the UAV's power equipment images, the improvements made to the original Faster-RCNN algorithm in this paper are as follows: First, to solve the problem of different resolutions and variables camera distances of images, we add a multi-dimensional feature extraction [8] module to adjust the network structure, combining top-down and bottom-up feature extraction methods to detect defects of different sizes in images of different scales. This improvement allows the algorithm to obtain a larger receptive field when paying attention to minor electrical equipment defects. Second, because the defective shapes of power equipment are primarily irregular, the traditional convolution shape is relatively single and does not have strong adaptability to defect detection of power equipment. We use deformable convolution [9] to enhance the feature extraction of defects of different shapes. Third, we use online hard example mining (OHEM) [10], and data augmentation [11] methods to solve the problem that there are generally only a single or a small number of electrical equipment defects in a single image captured by UAV. The improved algorithm framework pays attention to positive samples in the training process. It has a greater response to the area where the power equipment is located and where the power equipment is defective. More details are shown in Section III.

Our major innovations and contributions are as follows:

- This paper proposes an image defect detection method for power equipment based on deep learning. The method realizes integrated intelligent processing of defect detection, including UAV shooting, detecting and classifying defects automatically, and repairing electrical equipment.
- We add modules to optimize the performance of the native fault detection algorithm based on Faster-RCNN, including multi-scale feature extraction, deformable convolution, online hard

example mining, and data amplification.

- Through the 5847 images of UAV's electrical equipment with different resolutions for comparison experiments, it is proved that our algorithm for detecting defects of electrical equipment improves the accuracy by 60% on average compared with the benchmark algorithm. The method reduces the problem of excessive resource consumption in the maintenance of basic power facilities, and it is of great significance to national life and economic development.

This paper is organized as follows. In Section II we discuss some related work. In Section III, We will present our main idea. In Section IV we will show our experiment results, and we conclude in Section V.

2 Related Work

Although UAVs can capture numerous images and videos for electric equipment, they can not detect the defects of the equipment by automatically surveying those data. Therefore, image post-processing plays a critical role in defect detection after UVA image acquisition. The conventional processing method relies on the experience of the workers, which includes processing the images manually and marking the potential danger information via visual inspection. The cost of this approach is relatively high due to its experience dependence, labor enrichment, and material resource dependence. Thus, there is an unmet need to develop novel approaches to detect defects with better efficiency and lower cost.

Infrared image-based processing is one of the most commonly used methods for fault detection. Kazuo Yamamoto et al.[12] detected the transmission lines' faults in power equipment by fusing infrared images with RGB images, which enlarged the difference between those two images and contributed to targeting extraction. Based on this research, Walter Fetter Lages et al.[13] developed a real-time fault detection module for transmission lines. Although this method simultaneously provides and visualizes the faults information in power lines, it can only detect transmission lines with temperature difference characteristics. Moreover, the detection ability for some other types of faults is weak. Additionally, the operating ambient temperature of UAVs is relatively complicated. During the shooting process, the imaging attitude is uncertain, and there are many interferences. Furthermore, different infrared devices cannot work together, which limits the accuracy of the detection.

With the development of modern digital image technology, artificial intelligence, and other disciplines, diverse methodologies have been developed to optimize image post-processing and achieve varying results [14]. Since traditional image processing methods above-mentioned cannot address the unmet need in fault detection for power equipment, many scholars turn their attention to deep learning methods to explore novel detection methods with better efficiency. With the breakthrough of artificial neural network theory, an artificial neural network has gradually become an accurate and efficient solution to learn specific tasks and complete the corresponding tasks.

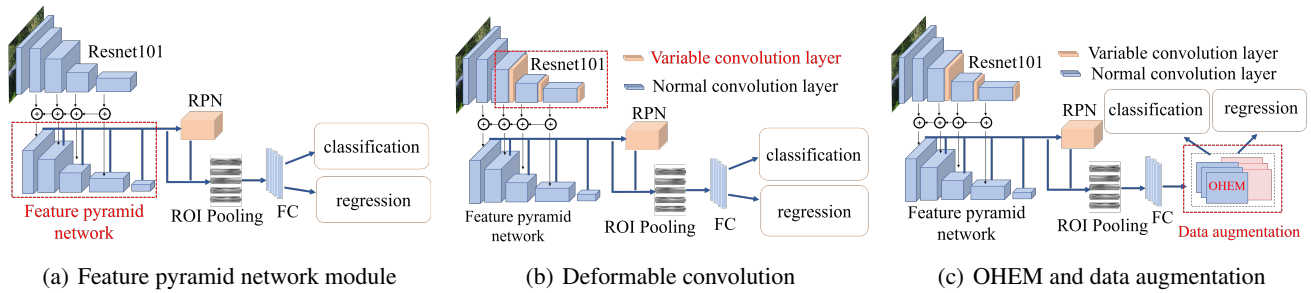


Figure 3: Schematic diagram of improvement

The literature [15] applied the AlexNet as a feature extraction network when using the random forest algorithm to detect insulator failures. Also, for detecting the fault in insulators, The literature [16] introduced multi-layer perceptrons to extract the location-related features and local contour features of the captured images. Moreover, The literature [17] adopts the neural network method to filter out the background noise in detecting the transmission line. Notably, with the combination of convolutional neural network and wavelet transform, The literature [5] can detect faulty transmission lines more accurately due to the better classification and extraction of fault features. However, the deep learning method has made progress in detecting electrical equipment with drone aerial photography, especially in accuracy and automation. The types of faults detected by related studies are relatively single. Most of them are in the preliminary exploration stage. Unlike the above deep learning methods, we will describe how to solve the problem more universally by improving the benchmark algorithm in the next part.

3 The proposed method

This section will introduce how to apply deep learning methods to identify and classify defects in images. This paper aims to realize intelligent processing as much as possible. Moreover, the Faster-RCNN algorithm is widely used in aerial inspection [18] [15] and has high stability [19]. Therefore, we select the Faster-RCNN as the benchmark framework for defect detection in UAV images of power equipment. However, defects are small targets, and the traditional convolutional layers decrease sensitivity when deepening in the actual scene, it is easy to ignore small targets. So we select the residual network [20] Resnet101 as the feature extraction network. As the overall accuracy of the benchmark algorithm is still not high, we consider the specific problems encountered in UAV photography to make three improvements to achieve higher accuracy in the following part.

3.1 Feature pyramid network

The detected objects in the images have various scales due to the different shooting distances and defect sizes. So the algorithm needs to detect targets of various scales.

Traditional deep feature extraction networks in the shallow layer can extract more detailed features that contain rich information. They also have a more incredible response to small-scale targets. Nevertheless, contextual semantic information misses due to the lack of receptive fields. The network continues to downsample while the feature extraction network is deepening. This way increases the receptive field corresponding to each area in the feature map. The features extracted at this time contain rich semantic information. However, the feature extraction networks have a more excellent response to large-scale targets while many details in the image are lost simultaneously. This shortcoming is fatal for defects detection of power equipment because power equipment defects are generally

small targets. Suppose we can extract multi-dimensional feature information and can achieve multi-scale feature extraction. In that case, the accuracy of detecting defects in electrical equipment captured by UAV will be further improved. So this paper implements a feature extraction process by extracting multi-scale features from the bottom-up and fusing the pyramid structure of different stages from the top-down to achieve the first stage optimization.

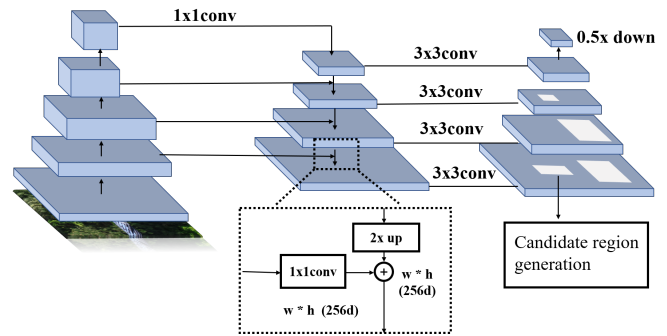


Figure 4: The process of multi-scale feature extraction using feature pyramid network, which integrates the top-down and bottom-down feature extraction methods

Fig. 4 shows the process of feature pyramid network. Firstly, the feature maps obtained in each stage are fixed as 1x1 convolution, then the number of channels is fixed as 256. Secondly, the lower-layer feature map is up-sampled by two times so that the feature map of the upper layer after convolution can scale the same size as the feature map of the lower layer. Then the bottom layer feature map fuses with the up-sampled upper layer feature through convolution and addition so that the high-level semantic information is merged with the detailed low-level information. Finally, The networks use 3x3 convolution to convolve the added feature maps to eliminate aliasing effects. Fig.3(a) shows the framework with the feature pyramid network.

3.2 Deformable convolution

The convolutional neural network proposes candidate regions and generates candidate frames by autonomously learning after the Faster-RCNN algorithm extracts image features. The corrosion of hardware is a common type of electrical equipment defect with irregular shapes. Additionally, The convolution kernel of traditional convolutional neural networks is generally rectangular or square. So it is a great challenge to detect the defects of the conductive line using the convolution with a regular structure. A fixed-shaped convolution kernel can sample a fixed-shaped area on the entire image, but the area covered by its receptive field is fixed for a given convolution kernel, which is not a good design. Moreover, The function of the high-level convolutional layer is to encode low-level features

and extract high-level features. However, targets at different locations often have different shapes and scales, limiting the coverage of the receptive field area, thereby restricting the detection of targets of different shapes. In the end, we adopt deformable convolution like Fig. 5 to accept defects. The convolution adjusts the scale and the coverage of the receptive field adaptively through learning, which will improve detection accuracy.

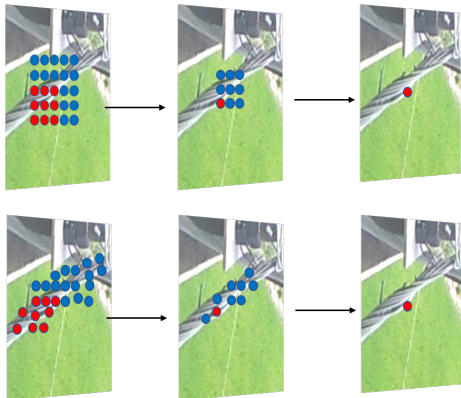


Figure 5: Comparison of the process between deformable convolution and traditional convolution

The conventional feature map convolution is usually divided into two steps in target detection:(1) sampling using a regular grid R over the input feature map x ; (2) summation of sampled values weighted by w . The final result is the value corresponding to the sampling position of the input feature map on the output feature map. In each position p_0 of the output feature map y , there is Eq.(1),

$$y(p_0) = \sum_{p_n \in R} w(p_n) * x(p_n + p_0) \quad (1)$$

where is an enumeration of each position of R . In the variable convolution, the regular grid R is expanded by the offset($\Delta p_n | n = 1 \dots N$), where $N = |R|$, the above formula is further transformed into Eq.(2).

$$y(p_0) = \sum_{p_n \in R} w(p_n) * x(p_n + p_0 + \Delta p_n) \quad (2)$$

After this change, sampling is no longer limited to the area corresponding to the regular grid R , but the offset forms an irregular area. As shown in Fig.6, the modified network predicts the offsets of each convolution position by the learned information before the convolution operation. After the position of the convolution changed by the offset, the network can pick the most suitable place to extract the features. The convolved area can be concentrated on defects as much as possible.

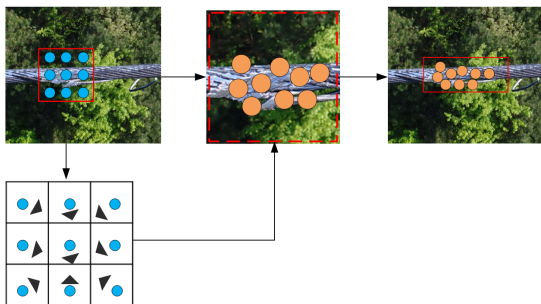


Figure 6: The process of deformable convolution to locate defects

As shown in Fig.3(b), this paper uses deformable convolution in the last three stages of Resnet101, enabling the feature extraction network to convolve appropriate locations at different scales. It improves the accuracy of small target detection.

3.3 Online hard example mining and data augmentation

There are often few or even no defects in the images of electrical equipment, which means that most of the candidate frames are based on the background. Thus, it is essential to correctly distinguish foreground and background. Correctly dividing the categories of positive samples for defect classification is also essential. This paper uses online hard example mining and data augmentation to improve the response-ability of the detection framework to targets.

Online hard example mining

In Faster-RCNN, the candidate region generation networks generate many candidate regions, many of which are negative sample boxes that do not intersect or intersect with the target box to be detected. These samples are *easy to train* samples. It is easy to reduce the loss function of the network by identifying these candidate sample frames as background. The ratio of the intersection with the foreground target frame (that is, an area ratio, which comes from the two candidate frames generated in the feature extraction stage, and divides their intersecting part and their merged part) of the *difficult to train* negative sample is relatively large. However, it does not exceed the set positive sample threshold. Therefore, it is necessary to focus on training those negative sample frames with a higher intersection ratio to be detected.

Online hard example mining can find *difficult to train* samples during the training process and increase weight of these samples to the loss function. The traditional method of regional candidate network to propose regional candidate frames is to randomly select 512 candidate frames to calculate the loss function according to the ratio of foreground and background 1:3 after the regression network refines and classifies the candidate frames. Instead, the online hard example mining method first calculates the loss of all the candidate boxes after refinement and classification, sorts the loss from high to low, and selects a total of 512 positive samples in a ratio of 1:3, and uses these samples to train network. Furthermore, it can improve the training effect by using the trained samples to train the network again. Then, such a method will make the sample's relative characteristics more evident than the traditional method.

Data augmentation

Data flipping. UAVs often capture images of electrical equipment from different directions, but the defects of the electrical equipment are invariant. It means the defect information will still exist after flipping an image vertically or horizontally. Therefore, during the training process, we randomly flip the images and the device defect frame marked by the image, and the data volume finally increases. Simultaneously, it improves the robustness of the detector by using these modified images to train the detector and allows the detector to learn useful information from images in different directions and angles.

Data scaling. UAVs usually do not maintain the same height in the process of photographing electronic equipment, and they cannot guarantee the same route and height every time. Furthermore, the resolution of UAVs of different brands is also quite different. Therefore, the images are randomly scaled to simulate the real environment in the training process of the model. Then the model will have a greater response to power equipment defects with different scales. Fig.3(c) shows that the detection framework with online hard example mining and data augmentation is effective in classification and regression.

After making relevant improvements, the next part of the experiment will prove that the improved detection framework has achieved obvious results.

4 Experiments

In this part, the traditional Faster-RCNN algorithm will be used as the basic framework to experimentally verify the ability of the improved system framework to detect image defects of power equipment.

4.1 Dataset

This experiment's data set consists of 5847 images of power equipment captured by UAVs, collected by one of the largest power companies in China. There are several types of electrical equipment defects in these images, such as foreign matters in the tower, damage of the insulator, abrasion of wires, and corrosion of hardware, and their resolutions are different. The data is divided into the training set, the verification set, and the test set at a ratio of 8:1:1. We verify the validation set when the training phase is completed (all training data are trained once) in training the model. We save the best model on the validation set and apply it to the test set to record the final result.

4.2 Evaluation indicators

Precision and Recall are important indicators for model evaluation. A comprehensive evaluation of recall rate and precision will be meaningful for evaluating defect detection of power equipment. Simultaneously, we also add the Average Precision and the mean Average Precision as measurement indicators according to the accuracy requirements of power system defect detection. The Average Precision refers to the average value of the maximum precision under the condition of calculating different recall rates, shown in Eq.(3):

$$AP = \int_0^1 p(r) dr \quad (3)$$

where AP stands for the Average Precision, and r refers to the recall rate. $p(r)$ represents the highest precision when the recall rate is r . The mean Average Precision refers to the average accuracy of each category, shown in Eq.(4):

$$mAP = \frac{AP}{C} \quad (4)$$

where mAP is the average accuracy of multiple types, and C is the total number of categories of the target to be detected. It should be noted that when calculating AP and mAP , we use the prediction frame whose intersection ratio with the real target object's outer frame (The ratio of the area of the intersection of the two candidate frames generated in the feature extraction stage to the area of the merged part) is more significant than 0.5 as the average accuracy calculated by the positive sample.

4.3 Comparative experiments

This paper sets up four stages of comparative experiments to study whether the improvement of Faster-RCNN is useful. Faster-RCNN with Resnet101 as the feature extraction network is set as the first group of comparative experiments recorded as experiment 1. Because the improvement of the benchmark system framework is divided into three parts, three sets of experiments are set to compare with experiment 1 during the investigation. Experiment 2 is the result of adding multi-scale feature extraction to the benchmark framework. Experiment 3 is the result of adding variable convolution based on experiment 2, and experiment 4 is the result of adding

online difficult sample mining and data augmentation based on experiment 3. The intuitive comparison of detection accuracy can reflect the effectiveness of the improved target detection framework for power equipment images.

4.4 Experimental results

As shown in Tab. 1, experiment 2 increases the average accuracy of experiment 1 by 19.3% after adding multi-scale feature extraction. According to our analysis, the reason is that the resolution of images captured by UAVs is relatively large. Images entered into the network are often reduced by a certain percentage. The feature extraction network also downsamples the images by 32 times, which will cause the inherently relatively small defects to be significantly reduced or even lost during the down-sampling process. These small defect features can be well preserved and detected on the high-resolution feature map after adding multi-scale feature extraction.

	BEN	FPN	DC	OHEM/DA	mAP
1	✓				43
2	✓	✓			62.3
3	✓	✓	✓		68.1
4	✓	✓	✓	✓	69.3

Table 1: The impact of algorithm improvement at each stage on the mean Average Precision. In particular, BEN stands for the benchmark algorithm, FPN stands for feature pyramid network, DC stands for deformable convolution, OHEM/DA stands for online hard example mining and data augmentation.

Additionally, the average accuracy of experiment 3 increased by 5.8% after using deformable convolution because experiment 3 is more adaptable to the shape of power equipment after adding deformable convolution. Finally, the average accuracy of experiment 4 increased by 1.2% by adding online hard example mining. We analyze that the model becomes easier to converge during the training process and learns from more difficult samples to get a more robust classification and detection capabilities. So experiment 4 further improves the average accuracy.

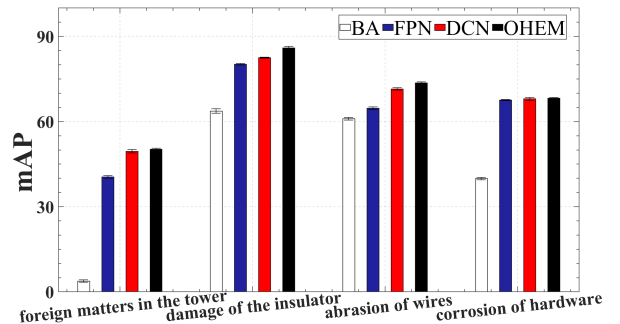


Figure 7: Detection accuracy of different algorithm

Fig. 7 shows the performance of different framework improvements in different categories of defects. In the original benchmark algorithm, this category can hardly be detected since the abnormal wire is small. With the addition of multi-scale feature extraction, the defects of this category are significantly increased. Because foreign matters in the tower are large and easy to detect, the benchmark target detection framework already achieve high detection accuracy, there are still some small-scale targets in this category of defects because the shooting distance of UAVs is relatively long, so multi-scale feature extraction can still improve the detection accuracy. De-

formable convolution can further improve accuracy when it detects frayed wires, damaged insulators, and foreign matters in the tower of different shapes. Finally, the algorithm is easy to jump out of the local minimum and converge better after adding online hard example mining.

	1	2	3	4	mAP
BA	3.8	63.7	64.7	39.9	43.2
OA	50.2	85.9	73.6	68.2	69.5
IR	12.210	0.349	0.138	0.709	0.609

Table 2: The accuracy improvement ratio of the improved optimal algorithm and the benchmark algorithm. 1,2,3,4 respectively represent abrasion of wires, foreign matters in the tower, damage of the insulator, and corrosion of hardware. BA,OA,IR respectively represent the benchmark algorithm, optimal algorithm, increase ratio. The results show that our algorithm exceeds the benchmark, which indicates the effectiveness of our approach.

Between comparison of the results of experiment 4 and experiment 1, it can be seen that the benchmark target detection framework is a better solution to the problem that minor defects in high-resolution power equipment images can hardly be detected. For foreign matters in the tower, damaged insulators, and large-scale clamp corrosion, the three major types of defects' accuracy are increased by about 30%, 15%, and 70% respectively. As shown in Tab 2, the average accuracy of each category has increased by about 60% eventually. However, the experimental results show that the accuracy of this method is still low for small-sized objects such as wires, so the results left much to be improved. Therefore, we can further try to use multi-task training methods to detect small-sized objects in a targeted manner while retaining the existing algorithm framework. Some results of our experiment are shown in Fig. 8.

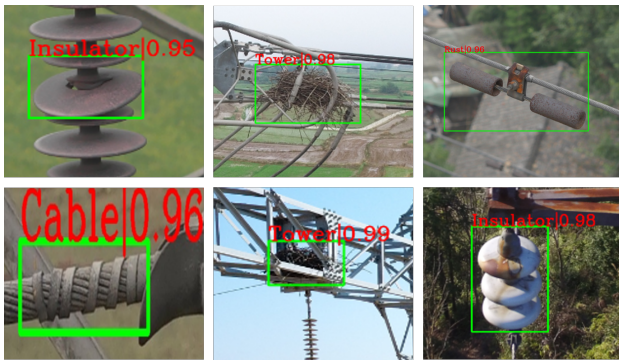


Figure 8: Visual display of defect detection: Tower refers to foreign objects in the tower, cable refers to abnormal conductors, insulator refers to damaged insulators, and rust refers to metal corrosion

5 Conclusion

In this study, we propose a detect deflection algorithm based on improved Faster-RCNN to process images captured by UAVs. We use a multi-scale feature extraction method to detect targets of different scales on images of different resolutions with high accuracy. Meanwhile, we find that traditional convolution cannot adapt flexibly due to the shapes of power equipment are multilateral. We introduce deformable convolution to solve this problem. Moreover, we improve the performance of defect recognition by studying the attention degree of the object to be detected in the labeled data. In general, this research has improved the accuracy of intelligent detection of defects in power equipment. Also, it has dramatically reduced the

consumption of human and material resources in the maintenance of basic power facilities, which has a high application value.

References

- [1] Ian Golightly and Dewi Jones. Corner detection and matching for visual tracking during power line inspection. *Image and Vision Computing*, 21(9):827–840, 2003.
- [2] Mohd Shawal Jadin and Soib Taib. Recent progress in diagnosing the reliability of electrical equipment by using infrared thermography. *Infrared Physics & Technology*, 55(4):236–245, 2012.
- [3] PN Druzhkov and VD Kustikova. A survey of deep learning methods and software tools for image classification and object detection. *Pattern Recognition and Image Analysis*, 26(1):9–15, 2016.
- [4] Xin Ma, Mingliang Li, Jinxi Kong, Siming Zhao, Wei Li, and Xiaohui Cui. Inspect characteristics of rice via machine learning method. *International Journal of Software Engineering and Knowledge Engineering*, 2020.
- [5] Weiguo Tong, Jinsha Yuan, and Baoshu Li. Application of image processing in patrol inspection of overhead transmission line by helicopter. *Power System Technology*, 34(12):204–208, 2010.
- [6] Alamuru Vani and Pessapaty Sree Rama Chandra Murthy. An adaptive neuro fuzzy inference system for fault detection in transformers by analyzing dissolved gases. In *2014 The 1st International Conference on Information Technology, Computer, and Electrical Engineering*, pages 328–333. IEEE, 2014.
- [7] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster r-cnn: Towards real-time object detection with region proposal networks. *arXiv preprint arXiv:1506.01497*, 2015.
- [8] Tsung-Yi Lin, Piotr Dollár, Ross Girshick, Kaiming He, Bharath Hariharan, and Serge Belongie. Feature pyramid networks for object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2117–2125, 2017.
- [9] Jifeng Dai, Haozhi Qi, Yuwen Xiong, Yi Li, Guodong Zhang, Han Hu, and Yichen Wei. Deformable convolutional networks. In *Proceedings of the IEEE international conference on computer vision*, pages 764–773, 2017.
- [10] Abhinav Shrivastava, Abhinav Gupta, and Ross Girshick. Training region-based object detectors with online hard example mining. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 761–769, 2016.
- [11] Luis Perez and Jason Wang. The effectiveness of data augmentation in image classification using deep learning. *arXiv preprint arXiv:1712.04621*, 2017.
- [12] Kazuo Yamamoto and Kimio Yamada. Analysis of the infrared images to detect power lines. In *TENCON'97 Brisbane-Australia. Proceedings of IEEE TENCON'97. IEEE Region 10 Annual Conference. Speech and Image Technologies for Computing and Telecommunications (Cat. No. 97CH36162)*, volume 1, pages 343–346. IEEE, 1997.
- [13] Walter Fetter Lages and Vinicius Scheeren. An embedded module for robotized inspection of power lines by using thermographic and visual images. In *2012 2nd International Conference on Applied Robotics for the Power Industry (CARPI)*, pages 58–63. IEEE, 2012.
- [14] Ting Yang, Liyuan Zhao, and Chengshan Wang. Review on application of artificial intelligence in power system and integrated energy system. *Automation of Electric Power Systems*, 43(1):2–14, 2019.
- [15] JF Li, QR Wang, and Min Li. Electric equipment image recognition based on deep learning and random forest. *High Voltage Engineering*, 43(11):3705–3711, 2017.
- [16] Yue Liu, Jianxiang Li, Wei Xu, and Mingyang Liu. A method on recognizing transmission line structure based on multi-level perception. In *International Conference on Image and Graphics*, pages 512–522. Springer, 2017.
- [17] S Hemamalini et al. Rational-dilation wavelet transform based torque estimation from acoustic signals for fault diagnosis in a three-phase induction motor. *IEEE Transactions on Industrial Informatics*, 15(6):3492–3501, 2018.
- [18] Q Chen, B Yan, R Ye, and XJ Zhou. Insulator detection and recognition of explosion fault based on convolutional neural networks. *J. Electron. Meas. Instrum*, 31:942–953, 2017.
- [19] Hong Zhang and Naiyan Wang. On the stability of video detection and tracking. *arXiv preprint arXiv:1611.06467*, 2016.
- [20] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.