

Hardware Detection Method of Transmission Line Patrol Inspection Image Based on Improved YOLOV4 Model

Wang Ning¹, Zhao Hanghang, Zheng Wulue, Wang Chaoshuo
CSG EHV Power Transmission Company, China

Abstract. In order to solve the problem of intelligent hardware detection in aerial images, a hardware target detection method based on improved YOLOV4 model is proposed. In order to solve the problems of dense hardware and occlusion in aerial images, the improved network based on channel and spatial hybrid attention mechanism can further improve the detection effect of dense occlusion hardware and reduce image false detection and missed detection. In order to solve the problem that there is a great error in the position of the detection frame caused by the interference between the hardware and the hardware and between the hardware and the background, the prior frame is optimized by K-means++, and it is determined that the anchors generated by K=12 is the best, and the detection boxes are more suitable for the target. The experimental results show that the proposed method solves the problems of missing detection, misdetection and inaccurate detection frame to some extent, in which the mAP (mean Average Precision) value of the performance index is increased from 65.03% to 70.72%. The research can lay a good foundation for further state detection and fault diagnosis of typical hardware.

Keywords. Hardware; YOLOV4; target detection; anchor optimization; mixed attention mechanism

1. Introduction

Transmission line fault is one of the important causes of large-scale power outages, and line operation and maintenance is an important link to ensure the safety of power system. Hardware is a metal component widely used in transmission lines[1]. High-precision automatic detection of hardware targets in aerial inspection images of transmission lines is the basis of its state detection and fault diagnosis.

The aerial hardware data set has the characteristics of diverse shapes of similar objects, diverse target angles, complex background and widespread target density. Even if the corresponding model parameters are adjusted and improved for YOLOV4[2], there are still two problems: (1). Due to a large number of shielding between the hardware and the hardware or between the hardware and the background, as well as the blurring of the features between the hardware and the background, the

¹ Corresponding Author, Wang Ning, CSG EHV Power Transmission Company, Guangzhou, China; E-mail: wangning1@ehv.csg.cn.

regression of the detection frame is not accurate. (2). For the dense part of the hardware, there will be the problem of misdetection and omission of small fittings.

Some common methods of hardware inspection are introduced in section 2, and their shortcomings are analyzed. In order to effectively solve the above two problems in section 3, this paper makes two improvements based on the YOLOV4 model. Aiming at the problems of dense hardware and occlusion in aerial images, an improved network based on channel and spatial hybrid attention mechanism is used in section 3.1 to improve the detection effect of the model on dense occlusion hardware and reduce image misdetection and missed detection. In order to solve the problem that there is a great error in the position of the detection frame caused by the interference between the hardware and the hardware and between the hardware and the background, the prior frame is optimized by using K-means++ in section 3.2, and it is determined that the anchors generated by K=12 is the best, and the detection frame is more suitable for the target to improve the accuracy of model detection. The experimental environment and data set information are given in section 4, and our model is analyzed quantitatively and qualitatively. The effect of the model is compared with two evaluation indicators, and then the test results are qualitatively analyzed. In section 5, the improved YOLOV4 is summarized.

2. Research status

With the increasing maturity of target recognition algorithm, the intelligent image processing technology of power system patrol inspection has also made great progress. At present, the existing transmission line image hardware detection methods are mainly divided into two categories: one is transmission line hardware recognition and detection based on traditional image processing and machine learning algorithm, the other is transmission line hardware recognition and detection based on deep learning algorithm [3]. The transmission line image recognition algorithms based on traditional image processing and machine learning algorithms mainly include: traditional power component recognition algorithms which rely on manually designed feature descriptors such as SIFT (scale invariant feature transformation), edge detector and HOG (directional gradient histogram)[4]. A step-by-step identification method of insulators based on target suggestion algorithm and structure search is proposed in reference [5]. Reference[6]after preprocessing the aerial image, the Haar features of the shock hammer are extracted from the image to train the AdaBoost classifier, and a method based on multi-view matching is proposed to reduce the missed detection rate of the shock hammer. The identification and detection of transmission line fittings based on deep learning algorithm mainly includes: in reference [7], FasterR-CNN[8]model is used in the detection of voltage sharing ring and shock hammer. In reference [9], a target decomposition and aggregation algorithm based on YOLOV3[10]model is proposed for insulator target detection in aerial images. As the latest deep learning target detection model based on regression method, YOLOV4 has high recognition accuracy and recognition speed. In view of the dense target set of transmission line hardware data, this paper improves the YOLOV4 model and achieves excellent results in 13 kinds of hardware detection.

3. Research methods

YOLOV4 is an object detection model which has been used in a variety of detectors since the release of YOLOV3, which can improve the detection accuracy of tricks, and improve on the basis of YOLOV3. Compared with YOLOV3 and YOLOV4, it not only ensures the speed, but also greatly improves the detection accuracy of the model. YOLOV4 has mainly made the following changes. For the backbone of feature extraction, they use CSPDarknet-53, compared to YOLOV3's Darknet, which uses cross-stage partial connection. The spatial pyramid pool [11] is used as the neck of the CSPDarknet-53 because it increases the receptive field, distinguishes the most important features, and does not slow down. The path aggregation network [12] is used instead of the feature pyramid network [13]. For the header, they used the original YOLOV3 network.

We recommend adding the following to YOLOV4, and we use the following methods to improve hardware detection on the hardware dataset.

3.1. Mixed Attention (MA)

Because the aerial image hardware detection is different from the existing general target detection, the existing classical general target detection mainly deals with images with a size of 1000×600 or less, while the size of the aerial images captured by the actual UAV patrol line is about 5000×3000 , of which more than 96% of the targets account for less than 0.05% of the images. Too small target proportion leads to a large number of missed detection. The use of multi-scale target detection in YOLOV4 can improve this problem. In the existing model, whether down-sampling the image or feature image to a fixed size will greatly lose the important information of the small target in the aerial image. In order to solve this problem, this paper uses the MA, to add MA to the down-sampling, which can make the model focus on the target feature more easily, so as to improve the detection rate of hardware.

First of all, the original attention mechanism of YOLOV4 is improved. On the basis of the spatial attention mechanism improved by the author, the information in the effective region can be more highlighted by adding the channel attention mechanism. The structure of the improved channel and space mixed attention mechanism is shown in figure 1. By adding the channel attention mechanism, the input features are averaged and maximized along the channel dimension. The output of the two is combined to get the feature descriptor. Finally, the convolution operation is used to encode, and the spatial attention map is obtained. The characteristic graph dimension of the output of the whole channel and spatial attention mechanism is the same as the input dimension, there is no need to make great changes to the network structure, and the parameters of the full connection layer are compressed according to the dimension reduction ratio of rust 16, which can weigh the balance between performance and propagation speed [14]. Due to the introduction of a more complex structure, the speed is slightly slower than the original spatial attention mechanism module, and the introduced computing time is mainly focused on the full connection layer and the pooled part of the channel. The full connection layer is used in the channel attention module, and the full connection layer is used because it has a global receptive field. in the channel attention module, what we need to get is the weight of each channel. This actually needs to include global information. The above improvements can not only help the network to classify and

locate objects more accurately, but also improve the problem of missing detection of small hardware under dense targets.

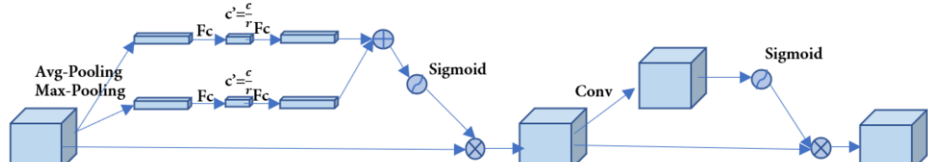


Figure 1. Based on channel and spatial attention mechanism.

3.2. Prior Anchor Optimization (PAO)

In the one-stage object detection algorithm, a priori anchor is widely used to set the initial dimension of the bounding box. YOLOV4 uses the K-means algorithm to generate nine prior anchor, from the coco dataset. In order to solve the defect caused by the random initialization of the clustering center in the original K-means algorithm, this paper introduces K-means++ to optimize the K clustering centers according to the following ideas: suppose n initial clustering centers have been selected ($0 < n < K$). When selecting the $n + 1$ cluster center, the farther away from the current n cluster center will have a higher probability of being selected as the $n + 1$ cluster center. The random method is also used when selecting the first clustering center ($n=1$). It can be said that this is also in line with our intuition: of course, the clustering centers are as far away from each other as possible. Although this improvement is intuitive and simple, it is very effective.

In order to determine the k value, we need to make a tradeoff between the ideal overlap value and the tolerable training time[15]. Increasing the k value will certainly increase the IoU(overlap)[16], but it will also greatly increase the computational overhead, because the number of convolution filters will increase linearly, thus increasing the training time. In the experiment, we draw the curve of the average IOU versus the K value, as shown in the figure, if the best K value is estimated according to the ElbowMethod method, if a certain K value makes the slope of the average IOU change obviously, then this K value is what we want. We find that the average IOU value of $K=12$ is 74.10% (while that of $K=9$ is 70.29%).

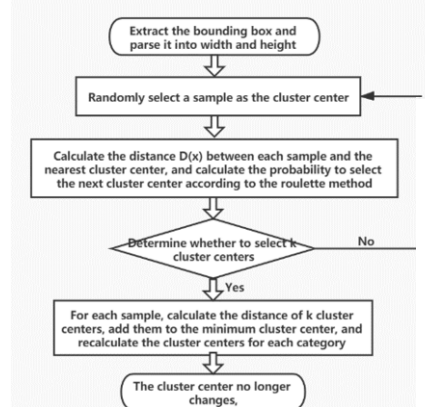


Figure 2. K-means++ Algorithm flowchart

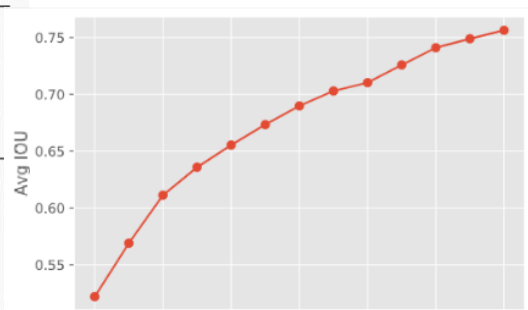


Figure 3. K-Avg IOU curve graph

4. Experimental results and analysis

The data set used in the first phase of this experiment includes 13 kinds of hardware targets, a total of 3774 images and 8738 tags, in which the ratio of training set to test set is 8:2. Most of the image sizes in the data set are 5000 × 4000 pixels, the input size of the GPU used is GTX1080, network is 416x416, the learning rate is set to 0.0013Maxbatchesbands 26000, the learning rate is adjusted in 20800 and 23400 iterations, and the weight is attenuated to 0.0005 using mosaic data enhancement.

In YOLOV4+MA, we use the improved channel and space hybrid attention mechanism, which can make the model better focus on the useful features in the image during training and detection, improve the problem of missed detection of dense targets, help to overcome the problem of image detection under different weather conditions, and improve the map@0.5 by 1 point.

In YOLOV4+PAO, we use IOU as a metric and Elbow Method to estimate the optimal K value of 12. Although the training time of the 12 prior anchor, is increased a little, the detection box generated by the optimized prior anchor, is more consistent with the real target, and the map@0.5 can be improved by 4 points.

With the comprehensive use of MA and PAO in YOLOV4, the experimental results are shown in Table 1, which can improve map@0.5 by 5.7 points.

Table 1. The original yolov4 and the method presented in this paper are used to detect the AP (Average Precision) values of each type of fittings and the mAP@0.5 values of 13 types of fittings in the transmission line aerial photograph data (YOLOV4 uses IoU > 0.5 by default).

Classes	YOLOV4	YOLOV4+MA+PAO
pre-twisted suspension clamp	45.36%	56.16%
bag-type suspension clamp	89.17%	91.81%
compression-type strain clamp	44.44%	45.24%
wedge-type strain clamp	48.46%	60.11%
hanging board	58.00%	61.17%
u-type hanging ring	70.48%	66.02%
yoke plate	55.49%	58.83%
parallel groove clamp	89.18%	92.14%
shockproof hammer	71.49%	77.19%
spacer	90.45%	89.75%
grading ring	23.21%	52.98%
shielded ring	92.99%	92.35%
adjusting board	66.71%	75.60%
map@0.5	65.03%	70.72%

In view of the fact that misdetection and omission have a great impact on the operation and maintenance of transmission lines, in order to meet the needs of hardware fault detection in actual power inspection, the practical effect of the combination of identification algorithm and company inspection work is realized. In addition to using the general evaluation index mAP in the field of target detection, this paper also uses the artificial intelligence algorithm evaluation method (val), which comprehensively measures the application effect of recall rate recall, accuracy precision and average accuracy AP, which is organized by the national network operation and inspection department. Its calculation formula is as follows:

$$val = 0.3 * recall + 0.2 * precision + 0.5mAP \tag{1}$$

It can be seen from Table 2 that the comprehensive use of PAO and MA, model has a higher recall rate, which can improve the problem of missing detection of hardware and improve the accuracy of detection, which is very important in the actual inspection task.

Table 2. Compare the original YOLOV4 and the method of this paper on the hardware data set for recall, precision, map and val of the test set.

Method	Recall	Precision	Map	Val
YOLOV4	0.69	0.76	0.65	0.693
YOLOV4+PAO+MA	0.74	0.8	0.707	0.736

In order to more intuitively show the high-precision detection performance and targeted improvement effect of this method, this paper increases the visualization experiment of the test data results. Figure 4 (a), (c) and (e) are the detection results of some images in the hardware image test set using only the adjusted original YOLOV4 model. Figure 4 (b), (d) and (f) are the detection results of the comprehensive application of PAO/MA method, including the target category, score and boundary box. As can be seen from figure 4 (a) and figure 4 (b), the light intensity of this sample is high due to the influence of weather. In this method, there is no missed detection, at the same time, the coincidence degree of the boundary box is higher, there is no serious deformation of the detection box, and too much redundant information is not classified into the detection frame. As can be seen from figure 4 (c) and figure 4 (d), for dense hardware detection, two small fittings in the original YOLOV4, u-type hanging ring and hanging board, have been missed, but all of them have been detected by this method. As can be seen in figure 4 (e) and figure 4(f), one of the lifting-type suspension clamps is disturbed by another heavy hammer fitting, which leads to false detection of fittings. The YOLOV4 model does not have a good generalization ability for this situation, but after using this method, the model can detect this interference more completely, indicating that it has better generalization performance.

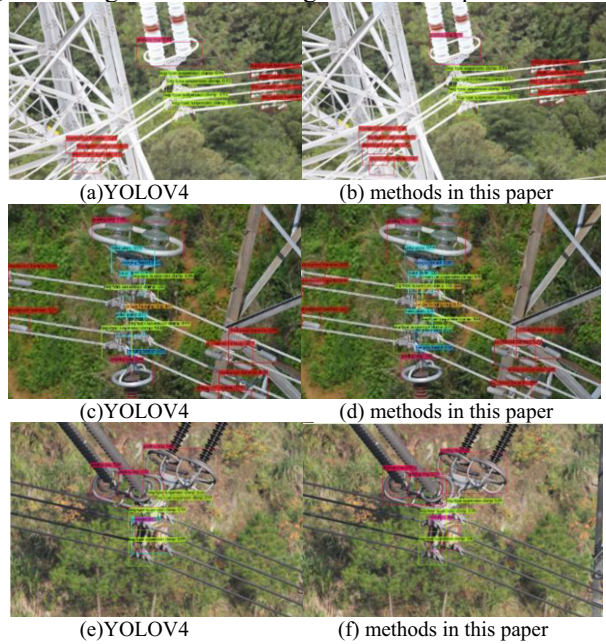


Figure 4. Visualization of YOLOV4 and methods in this paper

5. Conclusion

In this paper, a hardware detection method of transmission line aerial image based on improved YOLOV4 model is proposed. The improved mixed attention mechanism is used in YOLOV4 model, which can better detect small hardware with complex background and dense targets, optimize a priori anchor, and make the detection frame better match the real target frame. The improved detection method shows the excellent performance of small target and dense target detection under complex background, which improves the average accuracy of multi-targets to 5.69%, which lays a foundation for the subsequent defect state detection of hardware and its surface bolts. The method of this paper is quick to deploy and easy to operate, and can be extended to other application scenarios. In the future work, we will focus on solving the problem of model compression in order to deal with the embedded environment with limited space.

References

- [1] JIANG Xiuchen, LIU Yadong, FU Xiaofei, XU Peng, WANG Shaoqing, SHENG Gehao. Construction Ideas and Development Trends of Transmission and Distribution Equipment of the Ubiquitous Power Internet of Things[J]. High Voltage Engineering, 2019,45(5): 1345-1351.
- [2] Bochkovskiy, A., Wang C., Mark Liao, H. (2020). "YOLOv4: Optimal Speed and Accuracy of Object Detection". arXiv:2004.10934v1 [cs.CV] 2020, 23 Apr.
- [3] WANG Wanguo, TIAN Bing, LIU Yue, LIU Liang and LI Jianxiang. Study on the electrical devices detection in UAV images based on region based convolutional neural networks[J]. Journal of Geo-information Science, 2017, 19(2):256-263.
- [4] Zhang Xiaochun, Ouyang Guangze, He Hongying, Ding Yujie. Zero-insulator detection based on infrared images matching[J]. Electrical Measurement & Instrumentation, 2019,56(6):100-105.
- [5] Zhai Yongjie, Wang Di, Zhao Zhenbing. Recognition Method of Insulator Based on Object Proposals and Structure Research[J]. Journal of North China Electric Power University, 2016, (4):66-71.
- [6] Wang Sen. Research on Algorithms of Vibration Damper Detection on Power Line Image[D]. Beijing:Beijing Jiaotong University Masters's Degree Thesis, 2017.
- [7] Tang Yong, Han Jun, Wei Wen Li, Ding Jian, Peng Xin Jun. Research on part recognition and defect detection of transmissionline in deep learning[J]. Electronic Measurement Technology, 2018, 41(6): 60-65.
- [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, 2017, 39(6): 1137-1149
- [9] Gao Qiang, Lian Qiwang. Research on target detection of insulator in aerial image[J]. Electrical Measurement & Instrumentation, 2019, 56(5): 119-123.
- [10] YOLOV3: An Incremental Improvement. Dataset available from <https://pjreddie.com/media/files/papers/YOLOv3.pdf>.
- [11] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Spatial pyramid pooling in deep convolutional networks for visual recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 2015. 2, 4, 7, 37(9):1904–1916.
- [12] Shu Liu, Lu Qi, Haifang Qin, Jianping Shi, and Jiaya Jia., Path aggregation network for instance segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018, pages 8759–8768.
- [13] 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 (CVPR), 2017. 2, pages 2117–2125.
- [14] Woo, S., Park, J., Lee, J., & Kweon, I. CBAM: Convolutional Block Attention Module. (2018)ArXiv, abs/1807.06521.
- [15] Redmon, J., & Farhadi, A.. YOLO9000: better, faster, stronger. In Proceedings of the IEEE conference on computer vision and pattern recognition. 2017, pp. 7263-7271.
- [16] Zheng, Z.; Wang, P.; Liu, W.; Li, J.; Ye, R.; and Ren, D. (2020). Distance-IoU Loss: Faster and Better Learning for Bounding Box Regression. AAAI 2020