J. Meas. Sci. Instrum., 2024, 15(4): 571-578. http://jmsi.publish.founderss.cn/
DOI: 10.62756/jmsi.1674-8042.2024056 jmsi@nuc.edu.cn

Catenary dropper fault identification based on improved FCOS algorithm

GU Guimei[1*], WEN Bokang[1], JIA Yaohua[1], ZHANG Cunjun[2]

1. School of automation and electrical engineering, Lanzhou Jiaotong University, Lanzhou 730070, China;
2. China Railway Lanzhou Group Co., Ltd., Lanzhou 730030, China

***Corresponding author:** GU Guimei (xbbjb4@mail.lzjtu.cn)

Received: October 9, 2023 **Revised:** November 18, 2023 **Accepted:** January 1, 2024

Abstract: The contact network dropper works in a harsh environment, and suffers from the impact effect of pantographs during running of trains, which may lead to faults such as slack and broken of the dropper wire and broken of the current-carrying ring. Due to the low intelligence and poor accuracy of the dropper fault detection network, an improved fully convolutional one-stage (FCOS) object detection network was proposed to improve the detection capability of the dropper condition. Firstly, by adjusting the parameter α in the network focus loss function, the problem of positive and negative sample imbalance in the network training process was eliminated. Secondly, the generalized intersection over union (GIoU) calculation was introduced to enhance the network's ability to recognize the relative spatial positions of the prediction box and the bounding box during the regression calculation. Finally, the improved network was used to detect the status of dropper pictures. The detection speed was 150 sheets per millisecond, and the MAP of different status detection was 0.951 2. Through the simulation comparison with other object detection networks, it was proved that the improved FCOS network had advantages in both detection time and accuracy, and could identify the state of dropper accurately.

Key words: catenary dropper; fully convolutional one-stage (FCOS) network; defect identification; generalized intersection over union (GIoU); focal loss

0 Introduction

With the development of electrified railway, high-speed railway has gradually become the first choice for people's daily travel, and the requirements for the safety and reliability of high-speed railway operation are also constantly improving. The contact network is an important source of energy for the electrified railroad, and the stable power supply of the contact network directly affects the stability of high-speed railroad operation. The dropper exists in the chain suspension contact network system, which connects the contact line and the load-bearing cable, takes the role of conducting the current in the system, and improves the electrical quality of the contact line. The dropper can also improve the elasticity of the contact line and reduce the looseness of the contact line in the middle of the span. Therefore, it is important to ensure the normal working condition of the dropper to maintain the stability of the railroad power supply system and ensure the safe operation of high-speed railway[1,2].

With the application of contact network suspension detection and monitoring devices (4C) and various sensors, contact network condition monitoring means have been greatly enriched[3]. However, the amount of contact network picture data taken by the 4C device is huge. The accuracy rate of intelligent detection algorithm used at present is low, and the false detection rate and missing rate of contact network state detection results are high. Therefore, manual review is still required, which makes it less efficient and difficult to meet the needs of secure operation of the contact network[4-6]. With the development of deep learning, its characteristics of fast image processing, pervasiveness, and learning ability are found and gradually applied to the fault detection of contact network images taken by 4C-based devices, and good results are achieved. Yu et al.[7] used the faster R-CNN based on DenseNet for state identification of contact network dropper. Hu et al.[8] proposed an improved YOLO to detect dropper defects by changing the size and scale of the anchor boxes through clustering and using dual-scale feature fusion to achieve the recognition of dropper defects in different environments. Bian et al.[9] adopted a modified capsule network combined with a CV model, and took

© The Author(s) 2024. Published by Publishing Center of North University of China. The articles published in this open access journal are distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

vectors as input to achieve fault identification of the dropper. Chen et al.[10] used fully convolutional one-stage (FCOS) network to detect the unstressed condition of the dropper wire and achieved good results.

Although these algorithms can realize the defect recognition of contact network components, there is still some room for improvement in terms of detection accuracy and detection speed. An improved droplet defect state recognition network of FCOS contact network was proposed. Firstly, the parameter α in the focal loss function was adjusted to solve the positive and negative sample imbalance problem in the training process of dropper state recognition network. Secondly, the *GIoU* calculation was introduced to enhance the network's ability to identify the relative spatial positions of the prediction box and the ground-truth box during the regression calculation. Finally, the proposed network was simulated and compared with other networks to verify the recognition ability of the four states, such as normal dropper, slack, broken dropper wire, and broken carrier ring, as well as to verify that the recognition ability of improved FCOS network for ambiguous samples.

1 Methodology

Deep learning networks contain one-stage object detectors and two-stage object detectors, which differs in that the one-stage object detectors extract features to generate detection bounding boxes and then directly perform classification and regression loss calculation, such as YOLO networks and retinanet networks[11,12]. The two-stage target detector preselects the candidate bounding boxes after feature extraction, and then performs regression and classification loss calculations, such as the faster R-CNN network[13,14]. As a full-convolution single-stage target detector, FCOS network adopts pixel-by-pixel prediction method for target recognition after removing anchor box in optimized network structure, which improves detection speed. At the same time, after using the feature pyramid network (FPN), focus loss, and centrality, the detection accuracy of the network is also higher than that of the two-stage target detector. The network structure is shown in Fig.1.

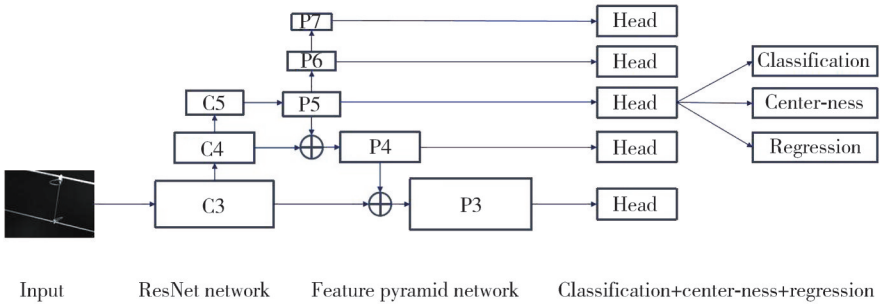


Fig. 1 Network architecture of FCOS

The features of the input images are extracted by the ResNet backbone network, and after deep and shallow feature fusion and hierarchical prediction through the FPN structure, the prediction results are output to the head layer for target classification and regression[15-17].

1.1 Anchor-free detector

The FCOS network is an anchor-free detector. The feature map is mapped back to the original image in a pixel-by-pixel manner, then positive and negative samples classification and regression are performed[18-20]. The pixel-by-pixel mapping means that $\boldsymbol{F}_i$ is represented as the feature map of layer i in the backbone network. When the position coordinate in the feature map is (x, y), it can be mapped back to the original image position through $\left(\left[\frac{s}{2}\right]+xs, \left[\frac{s}{2}\right]+ys\right)$, and s is the total step from the original map to the current feature map. When (x, y) is mapped into the ground truth box of the original image, it is considered as a positive sample, otherwise it is considered as a negative sample. By using the pixel-by-pixel position mapping instead of the anchor-based method for determining positive and negative samples, it avoids repeated prediction of the same pixel point and enables the FCOS network to correct the problem of a large number of candidate boxes and unbalanced positive and negative samples in the previous anchor-based one-stage networks. The FCOS network is trained in regression by the distance from the pixel location to the four boundaries of the bounding box, as shown in Fig.2.

Here $\boldsymbol{t^*} = (l^*, r^*, t^*, b^*)$, the 4D real vector represents the distance from the pixel location to the left, top, right, and bottom boundaries, respectively. During the regression, it can be described as[19]

$$l^* = x - x_0^i,\ r^* = x_1^i - x,$$
$$t^* = y - y_0^i,\ b^* = y_1^i - y, \tag{1}$$

where (x_0, y_0) and (x_1, y_1) denote the coordinates of the left-top and right-bottom position of the ground truth boxes, respectively.

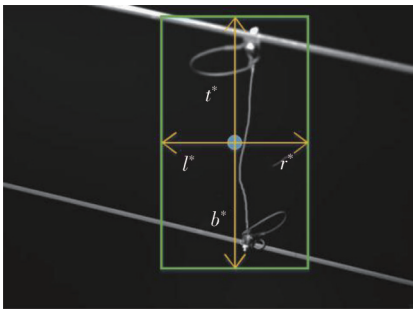


Fig. 2 4D Vector in FCOS Network

1.2 Feature pyramid network

When object recognition is carried out, if there are pixels in the image as shown in Fig. 3, after the feature map is mapped to the original image, these pixels fall in two overlapping boundary boxes, then the image is called fuzzy sample. The network performs the calculation of regression and classification on them due to the large-scale differences between two objects to be recognized in the same picture, which always results in the missed detection of small-scale targets. The FCOS enables the network to achieve more accurate detection of targets of different scales in the picture by adding FPN multi-level prediction.

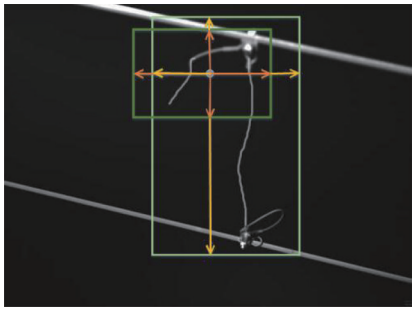


Fig. 3 Ambiguous sample

In the convolution operation of the network, shallow convolution extracts the detailed features of objects with definite information locations. In the deepening process of the network, although the features that can be extracted are more abundant, the location information of the object becomes fuzzy, which is easy to cause the loss of shallow information, thus reducing the recognition ability of the network for small objects. Therefore, FPN is introduced into the FCOS network structure, as shown in Fig. 1. FPN first performs a 1×1 convolution operation on C_5 of the ResNet network to get P_5. Then, the results of the 1×1 convolution operation of C_4 are combined with the 2×2 convolution upsampling operation of P_5 to obtain layer P_4. In the feature extraction stage, the same fusion method is used to obtain P_3 layer, which realizes the fusion of shallow layer and deep semantic information. The FCOS network also performs two successive convolutional down-sampling operations to obtain P_6 and P_7 layers, then the generated P_3 to P_7 layers are used to recognize objects at different scales, respectively. By setting different constraints on each layer, the layers of different scales can recognize objects of different sizes corresponding to the images, thus realizing the full utilization of each layer of FPN. The detection performance of the network is improved without basically increasing the computational effort of the original model.

1.3 Center-ness branch of FCOS network

In the pixel-by-pixel object prediction method of the FCOS network, many positive samples are far away from the center point and close to the target bounding boxes. In the regression calculation of the positions of such positive samples, the low-quality bounding boxes is produced, therefore, the center-ness branch is proposed in the classification calculation after performing FPN multi-level prediction. The corresponding value of center-ness branch is *Centerness**, and it is calculated according to the given (l^*, r^*, t^*, b^*) by[19]

$$Centerness^* = \sqrt{\frac{\min(l^*, r^*)}{\max(l^*, r^*)} \cdot \frac{\min(t^*, b^*)}{\max(t^*, b^*)}}. \tag{2}$$

The value range of *Centerness** is $(0, 1)$, indicating the deviation between the current calculated position and the center point. 1 is taken when the predicted point is located in the center of the image. After the corresponding center-ness value is calculated, it is multiplied with the output value of classification in the head layer, thus reducing the score far from the object center position and achieving the suppression of low-quality bounding boxes.

1.4 Loss function

1.4.1 Focal loss function

The focal loss function is expressed as[20]

$$FL(p_t) = -\alpha_t (1 - p_t)^{\gamma} \log(p_t), \tag{3}$$

where α_t is a weight parameter. When the sample is a positive sample, $\alpha_t = \alpha \in (0, 1)$ has no effect on its training weight; when the training sample is a negative sample, because its weight is higher, so that $\alpha_t = (1 - \alpha)$ makes the positive and negative sample of the

network close to balance. The P_t is expressed as[20]

$$p_t=\begin{cases}p, & y=1,\\ 1-p, & y=-1,\end{cases} \tag{4}$$

where $p\in[0,1]$ denotes the prediction of the targets by the network at the end of training; $y\in\{-1,1\}$ denotes the classification of the object labeled in ground-truth boxes, and $y=1$ when it is a positive sample.

When there is a positive sample object, and the network always misclassifies or predicts positive samples every time during the detection process, and every time is a low value, it is defined as a hard sample. The $(1-p_t)^\gamma$ in the function acts as a moderator to reduce the weight of the simple samples and focus the training more on the hard samples.

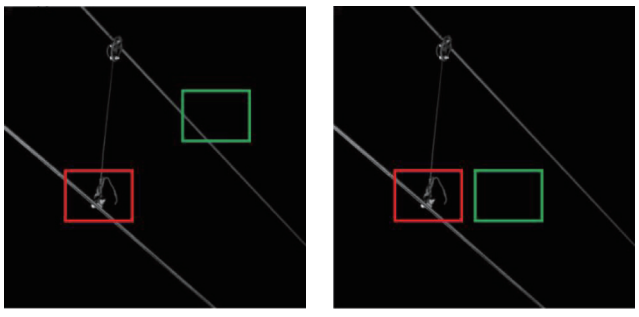
1.4.2 *GIoU*

Intersection over union (*IoU*) is used in object detection to determine the positive and negative samples, and the formula is[21]

$$IoU=\frac{|A\cap B|}{|A\cup B|}, \tag{5}$$

where A indicates the prediction boxes; B indicates the ground-truth boxes labeled in the input images; *IoU* indicates the ratio of the area of the intersection between the ground-truth box and the prediction box to the area of the concatenation, which represents the degree of overlap between the two boxes. The higher the *IoU* value, the closer the prediction box is to the ground-truth box, and the better the detection effect.

However, when the *IoU* values are the same, as shown in Fig. 4 (The ground truth box is red and the prediction box is green), it is reasonable to select the prediction box in Fig. 4(b) for regression calculation. Because the two boxes are the closest relative to each other in space.



(a) *IoU* is 0 inspace (b) Another case when *IoU* is 0 inspace

Fig. 4 Two cases where *IoU* is 0 in different spatial locations

But in this case, *IoU* does not have the ability to distinguish relative position in space. Therefore, the *GIoU* function is introduced for the FCOS network, and the expression is[21]

$$GIoU=IoU-\frac{C-(A\cup B)}{C}, \tag{6}$$

where C denotes the minimum area that can enclose the ground-truth box and the predicted box. When A and B overlap, $GIoU=1$. When A and B are far apart in space, *GIoU* tends to -1.

1.4.3 Loss gunction of FCOS network

The expression of the final loss function of the FCOS network consists of the classification loss function, regression loss function, and *Centerness**, the expression is[19]

$$L\left(\{\boldsymbol{p}_{x,y}\},\{\boldsymbol{t}_{x,y}\}\right)=\frac{1}{N_{\text{pos}}}\sum_{x,y}L_{\text{cls}}(\boldsymbol{p}_{x,y},\boldsymbol{p}^*_{x,y})+\frac{1}{N_{\text{pos}}}\sum_{x,y}IL_{\text{reg}}(\boldsymbol{t}_{x,y},\boldsymbol{t}^*_{x,y})+Centerness^*, \tag{7}$$

where $\boldsymbol{p}_{x,y}$, $\boldsymbol{t}_{x,y}$ denote the predicted values in classification and regression calculations; $\boldsymbol{p}^*_{x,y}$ and $\boldsymbol{t}^*_{x,y}$ denote the target values of the network during classification and regression in the train; N_{pos} denotes the number of positive samples; L_{cls} represents the classification focal loss; L_{reg} represents the *GIoU* loss in regression calculations; I is an indicator function, which is 1 when (x, y) is a positive sample, and equal to 0 otherwise.

2 Model training and result analysis

In order to demonstrate the effectiveness of the improved FCOS network model for fault detection of dropper, a series of simulation experiments were performed. All experiments were conducted in Pytorch deep learning framework using NVIDIA GeForce GTX 1650Ti for training.

2.1 Experimental dataset

In order to verify the identification ability of the contact network dropper based on the improved FCOS network, the dropper images taken at 4C were used as the data set, including 6 000 images of normal dropper, loose dropper line, broken dropper line, and broken current carrier ring. Among them, 5 400 were in the training set, and 600 were in the verification set. Before training, the training images in the datasets were firstly labeled with the corresponding label names zcdx, scdx, dldx, and zlhdl using LableImg software.

2.2 Training parameters setting

The initial learning rate of network training was set to 0.001. In the initial stage of training, the learning rate

was set to be larger to promote the updating of weights and the convergence of the network. However, with the increase of training times and learning rate, the loss function value will oscillate around the optimal value during parameter optimization. Therefore, the learning rate was set to decay as the number of training sessions increases. The FCOS network was trained a total of 90 000 times, the network learning rate decays to 1/10 of the current value at 60 000 and 80 000 times, which corresponds to 0.000 1 and 0.000 01, respectively. The larger learning rate in the first period is used to determine the optimal value region, and the decay of the learning rate in the later period facilitates the convergence to the optimal value.

2.3 Evaluation indicators

Precision rate (P), recall rate (R), average precision (AP), and mean average precision (MAP) are used as evaluation indicators, which is calculated by

$$P=\frac{TP}{TP+FP}, \tag{8}$$

$$R=\frac{TP}{TP+FN}, \tag{9}$$

where TP denotes the number of positive samples identified as positive by the network; FP denotes the number of negative samples incorrectly identified as positive; and FN denotes the number of positive samples incorrectly identified as negative. The AP value represents the area enclosed by the P-R curve of the network and the coordinate axis, and the MAP denotes the average of AP values of various classes.

2.4 Focal loss

In the object detection based on deep learning, the best training for hard samples is usually achieved when γ is set to the default value ($\gamma=2$). For the problem of imbalance between positive and negative samples in dropper state recognition, different detection networks take different values of α. When choosing the value, it is necessary to conduct experimental verification within its value range $(0,1)$, and then choose the value that makes the best detection effect. Table 1 shows a comparison of some of the data containing the optimal values.

When $\alpha=0$, it means that the optimization of the imbalance between positive and negative samples is removed from the focal loss function, and the detection result is the lowest. When $\alpha>0$, it can be concluded by comparison that the calculation effect using the focal loss function is generally improved than before. The detection is best when the α value is 0.1, so the parameters $\gamma=2$ and $\alpha=0.1$ in the loss function are finally determined.

Table 1 Comparison of network detection accuracy with different α values

α	Dropper state				MAP
	zcdx	scdx	dldx	zlhdl	
0	0.847 7	0.907 3	0.905 5	0.870 7	0.882 8
0.1	0.916 6	0.929 1	0.952 7	0.936 9	0.933 8
0.15	0.899 0	0.906 2	0.942 0	0.935 8	0.920 8
0.2	0.927 1	0.921 3	0.956 6	0.872 0	0.919 3
0.25	0.875 0	0.908 1	0.925 1	0.937 0	0.911 3
0.3	0.871 6	0.906 1	0.951 5	0.902 5	0.907 9
0.35	0.896 6	0.905 4	0.941 9	0.921 2	0.916 3
0.4	0.896 3	0.899 2	0.925 0	0.906 8	0.906 8

2.5 Optimized intersection and union ratio calculation

When calculating the intersection ratio between the ground truth box and the prediction box, the $GIoU$ that can describe the relative spatial position is used instead of the default IoU to optimize the network loss calculation. The backbone network was trained using ResNet50 network. Under the condition that other parameters of the network were the same, two calculation methods were used to train the dropper state recognition network. The results of network training are shown in Table 2.

Table 2 Comparison of dropper status recognition effect when training network with *GIoU* and *IoU*

Calculation method	Dropper state				MAP
	zcdx	scdx	dldx	zlhdl	
IoU	0.916 6	0.929 1	0.952 7	0.936 9	0.933 8
GIoU	0.925 7	0.937 9	0.978 1	0.963 1	0.951 2

It can be seen that the recognition ability for all dropper states of the network is improved after the optimization of the intersection and union ratio calculation.

2.6 Training results and detection performance of FCOS network

2.6.1 Training loss curve of FCOS network

Fig. 5 shows the total loss curve, classification loss, regression loss, and center-ness loss in the training of the FCOS network.

The loss value starts to stabilize at 25 000 times, and the learning rate decays to 1/10 of the original one at 60 000 times. The network loss value starts to decrease again and returns to stability quickly. The learning rate decays again at 80 000 times with no significant fluctuation in the loss curve. It means that the network has reached saturation at this point. The final network is trained a total of 90 000 times and then outputs the network model.

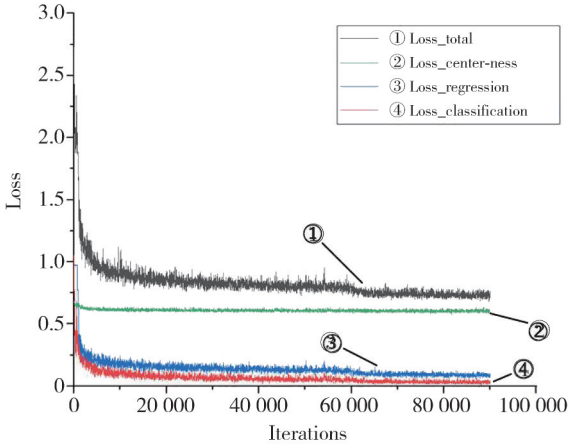


Fig. 5 Training loss curves of FCOS network

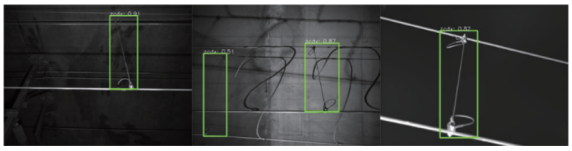
2.6.2 Comparison of effect of different networks

Using the same datasets, the detection performance of the improved FCOS network is compared with the two-stage network (such as faster R-CNN) and one-stage network (such as YOLO v4, RetinaNet), the data are shown in Table 3.

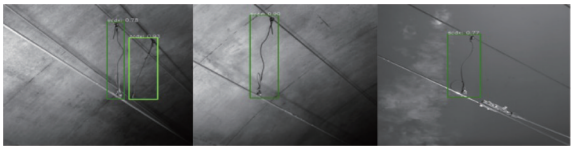
It can be shown that after optimizing the parameter α in the loss function and using the *GIoU* instead of *IoU* in the regression calculation, the recognition ability of the FCOS network for the dropper state is greatly improved. When the backbone network is selected as ResNet50, the improved FCOS network can achieve accurate recognition of the contact network dropper state while balancing the detection time and detection accuracy, and part of the detection effect is shown in Fig.6.

Table 3 Comparison of detection results of different networks

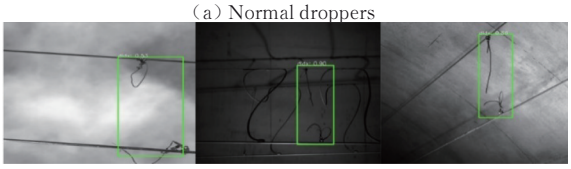
Network	*MAP*	Detection time/ms
Faster R-CNN	0.928 6	326
YOLO v4	0.857 7	225
RetinaNet	0.946 0	197
FCOS+Resnet50	0.951 2	150
FCOS+ResNet101	0.963 5	186



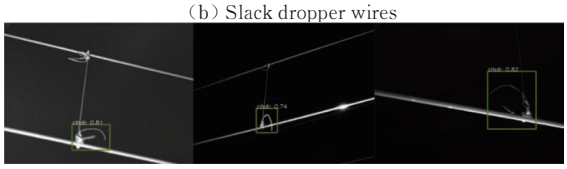
(a) Normal droppers



(b) Slack dropper wires



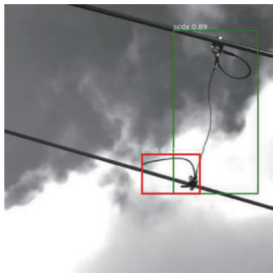
(c) Broken dropper wires



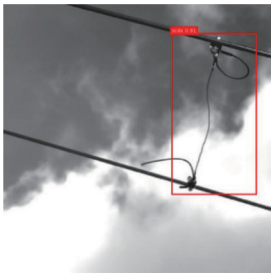
(d) Broken current-carrying rings

Fig. 6 Detection effects of improved FCOS network

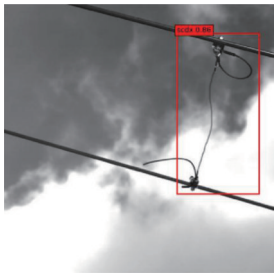
The comparison of the recognition effect for ambiguous sample between the improved FCOS network and other networks are shown in Fig.7.



(a) Improved FCOS



(b) RetinaNet



(c) Faster R-CNN

(d) YOLO v4

Fig. 7 Detection effects of different networks

As shown in Fig.7(a), when the object to be detected is an ambiguous sample, there are both dropper wire and broken current-carrying ring defects in the image.

The size difference of the two defective objects is large, and the overlapping area of the ground-truth boxes accounts for a large proportion of the small object, which often leads to the omission of small objects in the identification of fuzzy samples. The FCOS algorithm uses different feature layers for multi-level prediction of different scale objects through the FPN. At the same time, due to the overlap part being far from the center of the slack defect, the use of center-ness branch can reduce the contribution in its regression to the slack defect. By introducing GIoU, the network gives priority to the prediction box that is closer to the ground-truth box in spatial position for regression. So that the FCOS network can accurately identify the small-scale object in the ambiguous samples. Fig.7(b), (c), and (d) show the recognition effects of using other mainstream algorithms for ambiguous sample, which results in missed detection of small-scale object.

3 Conclusions

An improved FCOS dropper state recognition network was proposed. Since FCOS is a one-stage anchor-free detection algorithm, it has advantages over two-stage object detectors in terms of detection speed. Compared with the detector based on anchor point, the FCOS network eliminated the calculation of the area and scale of the anchor box, which further improved the detection speed of the network. And the focus loss function was used in the loss calculation, which solved the defects of the primary detector, such as the imbalance of positive and negative samples and the poor detection ability of hard samples, and improved the precision of target detection. Moreover, with the application of FPN multi-level prediction, *GIoU*, and center-ness branch in FCOS, it optimized the regression method of the prediction boxes and made the network have better detection performance for ambiguous samples than other networks. By comparing with other networks, it was proved that the improved FCOS network could achieve better results in identifying the states of dropper.

Acknowledgement

This work was supported by Natural Science Foundation of Gansu Province (No. 20JR10RA216).

Declaration of conflicting interests

The authors have no conflict of interests related to this publication.

References

[1] MEI X Y, GU G M, CHEN C. State identification of catenary clamp based on deep learning. Journal of Lanzhou Jiaotong University, 2022, 41(1): 61-67.

[2] LUO F L, YE W, WANG J. Defect detection of the puller bolt in high-speed railway catenary based on deep learning. Journal of Railway Science and Engineering, 2021, 18(3): 605-614.

[3] YU G W. Application of image processing based on hanging string defect detection algorithm in 4C system. Electrified Railway, 2021, 32(S1): 154-158.

[4] HUANG Y M, YUAN T C, YANG J. Research on identification method of catenary fault based on SVM. Computer Integrated Manufacturing Systems, 2018, 35(11): 145-152.

[5] HAN Y, LIU Z G, GENG X, et al. Feature detection of ear pieces in catenary support devices of high-speed railway based on HOG features and two-dimensional gabor transform. Journal of the China Railway Society, 2017, 39(2): 52-57.

[6] BAI R M. Detection and recognition of catenary string and pantograph skateboard based on image processing technology. Chengdu: Southwest Jiaotong University, 2017.

[7] YU X N, GU G M, WANG Y P, et al. Catenary dropper fault detection method based on faster R-CNN. Journal of Lanzhou Jiaotong University, 2021, 40(2): 58-65.

[8] HU D, JIN W D, TANG P. Detection of railway dropper based on double scale features. Electric Engineering, 2020, 15(20): 64-68.

[9] BIAN J P, HAO J X, ZHAO S, et al. Fault identification and location of catenary suspension based on improved capsule network. Transactions of China Electrotechnical Society, 2020, 35(24): 5187-5196.

[10] CHEN Q, PENG J S, YAN Y F, et al. Method based on FCOS and resNet50-FL for identifying stress-free Dropper. Journal of the China Railway Society, 2021, 43(10): 36-42.

[11] WANG Q H, GU W, CAI P Z, et al. Detection method of double side breakage of population cotton seed based on improved YOLO v4. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(1): 389-397.

[12] HOU Q Z, SUN J Y, WANG H, et al. Runway edge lights brightness detection based on improved RetinaNet. Laser & Optoelectronics Progress, 2022, 59(2): 192-200.

[13] YIN T P, YANG J. Detection of steel surface defect based on faster R-CNN and FPN//2021 7th International Conference on Computing and Artificial Intelligence, April 23-26, 2021, Tianjin, China. New York: ACM, 2021: 15-20.

[14] GU G M, CHEN C, YU X N, et al. Target location algorithm of contact network pipe cap based on improved Faster R-CNN. Laser & Optoelectronics Progress, 2022, 59(4): 140-150.

[15] ZHANG Z H, JIA W K, SHAO W J, et al. Green apple based on optimized FCOS in orchards. Spectroscopy and Spectral Analysis, 2022, 42(2): 647-653.

[16] TAN Y D, YU D, HU Y. An application of an improved FCOS algorithm in detection and recognition of industrial instruments. Procedia Computer Science, 2021, 183: 237-244.

[17] LIU S, CHI J N, WU C D. FCOS-Lite: an efficient anchor-free network for real-time object detection//2021 33rd Chinese Control and Decision Conference (CCDC), May 22-24, 2021, Kunming, China. New York: IEEE, 2021: 1519-1524.

[18] LONG Y, LI N N, GAO Y, et al. Apple fruit detection under natural condition using improved FCOS network. Transactions of the Chinese Society of Agricultural Engineering, 2021, 37(12): 307-313.

[19] TIAN Z, SHEN C H, CHEN H, et al. FCOS: fully convolutional one-stage object detection//2019 IEEE/CVF International Conference on Computer Vision (ICCV), October 27-November 2, 2019, Seoul, Korea

(South). New York: IEEE, 2019: 9626-9635.
[20] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(2): 318-327.
[21] REZATOFIGHI H, TSOI N, GWAK J, et al. Generalized intersection over union: a metric and a loss for bounding box regression//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. New York: IEEE, 2019: 658-666.

基于改进FCOS算法的接触网吊弦故障识别

顾桂梅[1*], 温柏康[1], 贾耀华[1], 张存俊[2]

1. 兰州交通大学 自动化与电气工程学院, 甘肃 兰州 730070;
2. 中国铁路兰州局集团有限公司, 甘肃 兰州 730030

摘　要: 接触网吊弦工作环境恶劣, 在列车行驶过程中还会遭受受电弓的冲击作用, 可能出现吊弦线松弛、断裂和载流环断裂等故障。由于吊弦的故障检测网络智能性低、准确率差, 本文提出一种改进的全卷积一阶段目标检测(Fully convolutional one-stage object detection, FCOS)网络来提高对吊弦状态的检测能力。首先, 通过调节网络焦点损失函数中的α参数消除网络训练过程中正负样本不平衡问题。其次, 引入广义交并比(Generalized intersection over union, GIoU)计算, 增强网络在回归计算时对预测框和目标框相对空间位置的识别能力。最终, 使用改进后的网络对吊弦图片进行状态检测, 检测速度为150张每毫秒, 对不同状态检测的*MAP*为0.951 2。通过与其他目标检测网络的仿真对比, 证明了改进后的FCOS网络在检测时间和精度上同时具有优势, 能准确地对吊弦状态进行识别。

关键词: 接触网吊弦; FCOS网络; 缺陷识别; 广义交并比(GIoU); 焦点损失

引用格式: GU Guimei, WEN Bokang, JIA Yaohua, et al. Catenary dropper fault identification based on improved FCOS algorithm. Journal of Measurement Science and Instrumentation, 2024, 15(4): 571-578.