

# Infrared Thermal Image Fault Detection Based On YOLOV3-L

Si Gao, Yiting Ruan, Qiancheng Hong, Decai Yin

(Hangzhou Hikvision Digital Technology Co., Ltd.)

## Abstract

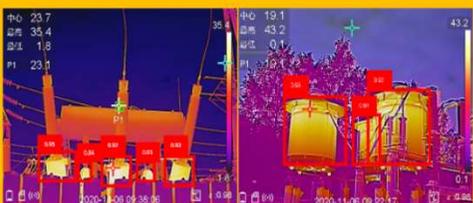
In the substation scenario, since the substation equipment is exposed to nature for a long time, it is easy to generate heat and cause the failure of the substation equipment. In order to effectively detect potential faults, a thermal infrared image detection method based on improved YOLOV3 (You Only Look Once) is proposed. This method is based on the overall framework of YOLOV3, and uses another lighter convolution structure to replace the standard convolution in Darknet-53, thereby reducing model parameters, improving the detection speed of thermal infrared images, and optimizing the use of The latter NMS algorithm optimizes the detection accuracy. The experimental results show that the proposed YOLOV3-L method has almost a 100% increase in FPS compared with YOLOV3, while the accuracy is almost unchanged, reaching a MAP value of 95.4%.

## EXPERIMENT



(a)

(b)



(c)

(d)

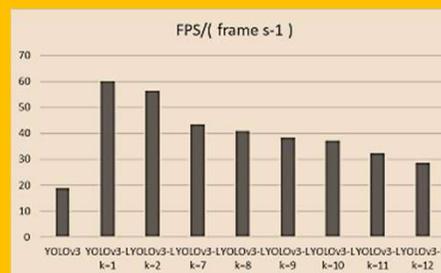
As shown in Figure, the proposed method herein has good detection performance, containing scale different regions in figure (a) and (b), but still can be detected, including multiple regions in (c) is not omitted, and (d) contains some overlapping regions, which are also satisfactorily detected under the method herein.

## Methods

Darknet53		LDarknet53	
conv32,3*3		conv32, 3*3	
conv64,3*3/2		conv64, 3*3/2	k=k1
conv32,1*1	*1	conv32,1*1	*1,
conv64,3*3		lconv64,3*3	k=k2
residual		residual	
conv128,3*3/2		conv128,3*3	
		/2	
conv64,1*1	*2	conv64,1*1	*2,
conv128,3*3		lconv128,3*3	k=k3
residual		residual	
conv256,3*3/2		conv256, 3*3	
		/2	
conv128,1*1	*8	conv128,1*1	*8,
conv256,3*3		lconv256,3*3	k=64
residual		residual	
conv512,3*3/2		conv512, 3*3	
		/2	
conv256,1*1	*8	conv256,1*1	*8
conv512,3*3		lconv512,3*3	k=128
residual		residual	
conv1024,3*3/2		conv1024, 3*3/2	
conv512,1*1	*4	conv512,1*1	*4
conv1024,3*3		conv1024,3*3	
residual		residual	
Avg pool		Avg pool	
dense		dense	

In the table , lconv2d, is used for 5 convolutional modules . For convolutional layers with many channels, too small k value causes the problem of feature loss, and more parameters when k is too large. If the number of output channels is 256 and 512, k takes 64 and 128, respectively, and the parameter selection for k1,k2,k3 is determined by experiment.

## Results



By analyzing the data in the two Figure , we can draw some conclusions as follows:

(1) When k=1, there is a case of feature loss. Although the speed is increased by more than two times, the accuracy is very low, and when k=2, the accuracy does not increase but decreases, which indicates that when k is small , the feature loss is severe, and a slight increase does not improve the situation.

(2) When k=10, the accuracy reaches 0.954. Compared with the original YOLOV3, the speed is almost doubled, but the accuracy is only reduced by 0.4%. Compared with k=11 and 12, when k=10 not only Higher accuracy and faster speed. This shows that in this scenario, k=10 is more appropriate.

(3) When k=11, 12, although the parameters are increased, the accuracy has not been improved, which indicates that the parameter value of the algorithm has a peak. At the same time, the algorithm can only improve the speed, but cannot submit the target detection. accuracy.

## Acknowledgment

We would like to thank the participants and companies who take part in this study. This work is supported by Research on multi-spectral spectral video intelligent inspection technology for substation equipment under Science and technology project of State Grid Corporation of China (5200-201919048A-0-0-00)

## Conclusion

Aiming at the problem of fault detection of power equipment, this paper proposes a method named YOLOV3-L based on the YOLOV3 algorithm. This method effectively reduces the parameters of the network by improving the standard convolution. By exploring the parameter k, this paper obtains The best result is obtained when the parameter k=10 is used. Through visualization experiments.