

PAPER • OPEN ACCESS

Critical fault diagnosis method of small components of power transmission based on optimizational YOLOv3 algorithm

To cite this article: Jia Chen Chen 2020 *J. Phys.: Conf. Ser.* **1619** 012001

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



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

239th ECS Meeting with IMCS18

DIGITAL MEETING • May 30-June 3, 2021

Live events daily • Free to register

Register now!

Critical fault diagnosis method of small components of power transmission based on optimizational YOLOv3 algorithm

Jia Chen Chen¹

¹ Chien-Shiung Wu College, Southeast University, Nanjing 210000, China

¹ 894717825@qq.com

Abstract. In order to solve the problem that the image recognition algorithm in UAV (Unmanned Aerial Vehicle) inspection of power transmission has low recognition rate for small, indiscoverable and high-risk component failures, this paper proposes a critical fault diagnosis method based on the optimizational YOLOv3 algorithm. We add SPP net under the basic framework of the original YOLOv3 and extract multi-scale depth features in the same convolutional layer which can transform all the input images into a uniform size and solve the problem caused by different input image sizes and formats. An optimizational scale prediction framework of YOLOv3 is proposed, which expands the original three-scale prediction frameworks to four scales and carries out multi-scale integration. This improvement is used to strengthen the feature extraction of small and indiscoverable fault. The optimizational algorithm is verified by using the dataset of electrical components of transmission system in the actual scene which have small and indiscoverable faults. The results show that the optimizational YOLOv3 improves the recognition rate of the selected categories by 3.98% on average, and increases the recognition rate of the small size and indiscoverable fault by 11.95% on average.

1. Introduction

With the continuous development of China's smart grid, the technology of using UAV to take pictures of power transmission and transmit them back to the background server for component fault identification is gradually applied to the daily maintenance of power system on more occasions. However, with the practical application, this technology has exposed many problems. For large-scale electrical components, such as insulators, vibration dampers, the recognition rate is high when using artificial intelligence image recognition technology to detect defects and faults; while for small-scale electrical components, such as U-bolts, strain clamps, it is likely to miss-detection and false-detection. And such undetectable defects or faults often have a higher risk to cause damage to the transmission system. Therefore, it is of great significance for the development of patrol inspection technology to study the image recognition algorithm which can increase the defect recognition rate of electrical components which are difficult to find and small in size.

At present, image recognition algorithms can be divided into two categories: algorithms based on region proposal method and regression method. The former mainly includes Fast R-CNN, Mask R-CNN, etc. The training model of the latter has the advantages of high detection speed, and the more typical ones are YOLO series. In order to adapt to different engineering application scenes, many improved algorithms are proposed. The GIOU calculation method for the optimization of helmet detection is proposed. It combines with the goal function of the YOLOv3 algorithm to form a new goal function, so that the goal function local optimum equals to IOU local optimum[1]. An algorithm for adding a



convolutional layer module to the YOLOv3 network structure and roughly adjusting the size of the ground anchor frame on the feature map is proposed[2]. Another optimization algorithm for evaluating air quality has been proposed, which reassembles the YOLOv3 network to include two modules, the multi-anchor detection mechanism and the convolution voting network[3]. An improved Faster-RCNN model for patrol image multi-morphology is presented. It generates several target regions through the region recommendation strategy network, then trains the convolutional neural network based on the actual patrol image sample dataset to improve the learning effect of parameters, and finally optimizes the parameter weights by using the regularization method to improve the detection speed[4]. A new bearing fault detection for three-phase induction motors using Artificial Neural Network (ANN) is proposed[5]. There is a paper about pattern recognition using Bidimensional Empirical Mode Decomposition (BEMD) for mammogram images to extract five statistical textural features and Artificial Neural Network (ANN) to distinguish mass and non mass tissue based on these features[6].

The existing improved algorithms mostly aim to improve the image recognition accuracy. Problems such as small parts in transmission line and tower that are difficult to recognize still exist, and current researches on small target recognition are also incomplete. An image target detection algorithm for power equipment in smart substation using single shot multi-box detector (SSD) is presented. It is optimization by adding feature extraction layer according to the related features of power equipment dataset in smart substation, redesigning the number and proportion of feature prediction boxes, and using soft punishment non-maximum suppression (Soft-PNMS) and other improved methods. However, this method only involves substation equipment, and does not consider smaller power components in the transmission system, so it is relatively difficult to identify[7]. A single-stage deep neural network (DF-YOLOv3) algorithm is proposed. It can accurately locate and classify relatively small targets in complex scenes. The convolutional feature maps of different sizes are designed in the algorithm, and then integrated with the corresponding scales in the residual network to form a feature pyramid. This method is more accurate for small target detection, but it still has the drawbacks of complex calculation steps and long time-consuming[8].

Therefore, in order to solve the problem that the current pattern recognition algorithms detect small size electrical components with the high rate of miss and wrong detection, this paper presents an optimization YOLOv3 deep learning algorithm based on regression. We improve the YOLOv3 based on the framework of Darknet, adds several SPP net after the specific convolution layer, and optimizes the multi-scale prediction framework of YOLOv3. After the frame is modified, the recognition rate of the model for the smaller electrical components is improved.

2. About YOLOv3 and some related technologies

2.1. YOLOv3 and its multi-scale prediction framework

The structure of YOLOv3 is composed of two parts: the convolutional neural network and the prediction network of Darknet-53. The structure of Darknet-53 is used to extract features. Figure.1 shows the structure of Darknet-53, which uses continuous convolutional layers (3×3 and 1×1) and a large number of residual skip connections. The prediction network has three multi-scale prediction frameworks, each of which predicts three anchor boxes, so that nine anchor box (10×13 ; 16×30 ; 33×23 ; 30×61 ; 62×45 ; 59×119 ; 116×90 ; 156×198 ; 373×326) can be obtained, and they can be allocated to three characteristic graphs of 13×13 , 26×26 and 52×52 according to their size.

	Type	Filters	Size	Output
1 ×	Convolutional	32	3 × 3	416 × 416
	Convolutional	64	3 × 3 / 2	208 × 208
	Convolutional	32	1 × 1	
	Residual	64	3 × 3	208 × 208
2 ×	Convolutional	128	3 × 3 / 2	104 × 104
	Convolutional	64	1 × 1	
	Convolutional	128	3 × 3	
	Residual			104 × 104
8 ×	Convolutional	256	3 × 3 / 2	52 × 52
	Convolutional	128	1 × 1	
	Convolutional	256	3 × 3	
	Residual			52 × 52
8 ×	Convolutional	512	3 × 3 / 2	26 × 26
	Convolutional	256	1 × 1	
	Convolutional	512	3 × 3	
	Residual			26 × 26
4 ×	Convolutional	1024	3 × 3 / 2	13 × 13
	Convolutional	512	1 × 1	
	Convolutional	1024	3 × 3	
	Residual			13 × 13
	Avgpool		Global	
	Connected		1000	
	Softmax			

Figure 1. The structure of Darknet-53

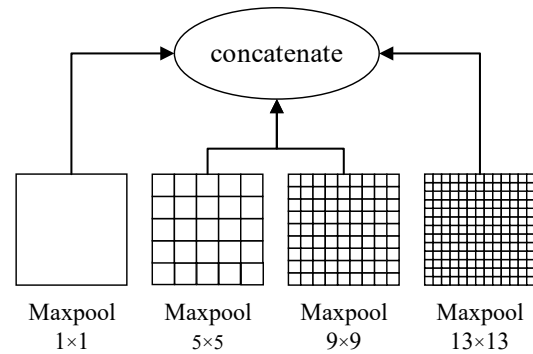


Figure 2. The structure of SPP net

2.2. SPP net

SPP net uses the spatial pyramid pooling model [9] to divide the image from the fine space to the rough space, and gather its local features, which can be used to enhance the detection efficiency of small-scale objects. The structure of SPP net is shown in figure.2. The SPP net used in this paper consists of four parallel kernel sizes maximum pool layers(1×1 , 5×5 , 9×9 , 13×13).

3. The improvement of YOLOv3 network structure

3.1. Add SPP net behind the specific convolution layer

In the process of UAV inspection of transmission system, the images of small components of it taken by UAV are often blurred, which makes the image recognition algorithm difficult to recognize. If the image is cut and deformed according to the original YOLOv3 network, it is likely that critical information will be lost. In this condition, such small size and difficult to detect defects will seriously affect the safety of the power system. Aiming at this problem, the YOLOv3-SPP network[10] used in this algorithm improves the YOLOv3 network in convolutional calculation. Under the premise of retaining the basic framework of the YOLOv3 network, the SPP net is integrated between the fifth layers and the sixth layers of each test header. Its working principle is shown in figure 3. The YOLOv3-SPP network obtains multi-scale features in the same layer through the SPP net, so as to more efficiently obtain the image information of the small-size power components in the face of the different sizes of the power equipment in the picture. At the same time, the SPP net uses the pooling operation of fixed block to realize the output of the same size for the input of different sizes, which is conducive to solving the problem of different sizes of pictures taken by UAV. For example, for the input size of 512×512 feature map, through the SPP net, the feature map is transformed into a matrix of $(169 + 81 + 25 + 1) \times 512$, which is input into the later 1×1 convolutional layer and expanded into one-dimensional matrix to output a fixed size eigenvector. For the coordinate mapping transformation of SPP net, the following formula is used:

$$x' = \left\lfloor \frac{x}{S} \right\rfloor \pm 1 \quad (1)$$

$$y' = \left\lfloor \frac{y}{S} \right\rfloor \pm 1 \quad (2)$$

where x and y are the horizontal and vertical coordinates of the coordinate before mapping; x' , y' are the

horizontal and vertical coordinates of the coordinate after mapping; S is the product of all strides.

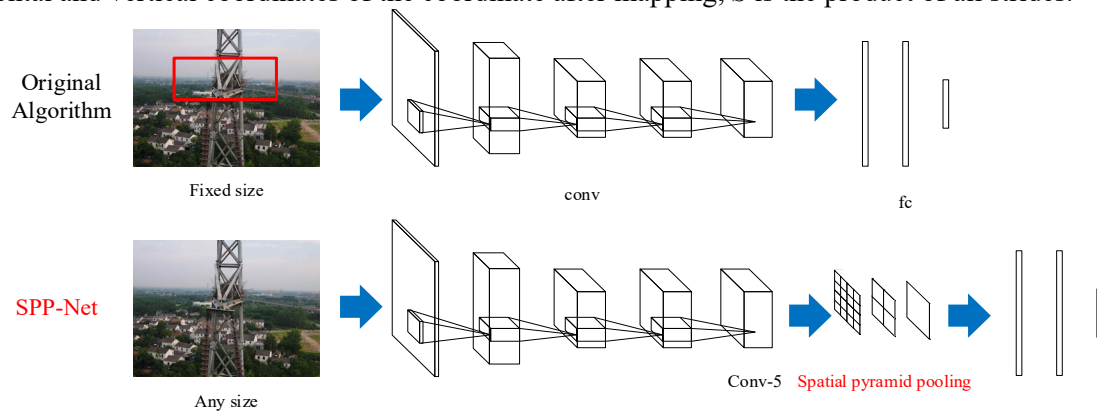


Figure 3. The flow diagram of original YOLOv3 and YOLOv3 with SPP net

3.2. Improvement of multi-scale prediction framework

In view of the large variety of small power equipment in transmission system and the difficulty to find their fault defects, this paper modifies the original three-scale prediction frameworks in YOLOv3, and then uses the idea of FPN (feature pyramid networks) to integrate the modified scale prediction framework. Experiments show that in the neural network, the shallow network pays more attention to the low-level detail information of the image, and the deeper layer of the network pays more attention to the high-level semantic information of the image. Therefore, if we neglect the low-level image feature information taken from the shallow network, it is very easy to leak or mistake detection of the defect of smaller size components. After integration, the network not only preserves high level image feature information, but also obtains more low-level image feature information to detect smaller targets.

The working principle of FPN is shown in figure.4. The corresponding branches of layers 1, 2 and 3 are the bottom-up pathway, which is used for pre-training. It is to select a picture to be processed for pre-processing, and then send it to the pre-training feature network (ResNet, Residual Network). Layers 4, 5 and 6 correspond to the top-down pathway, which is the core of the FPN algorithm. The network integrates the information of multiple feature layers, and the information obtained from each feature layer is reused by the next layer, which is compatible with low-level and high-level feature information. In the following part of 3.3, the method to determine the specific number of scale prediction framework will be introduced in detail.

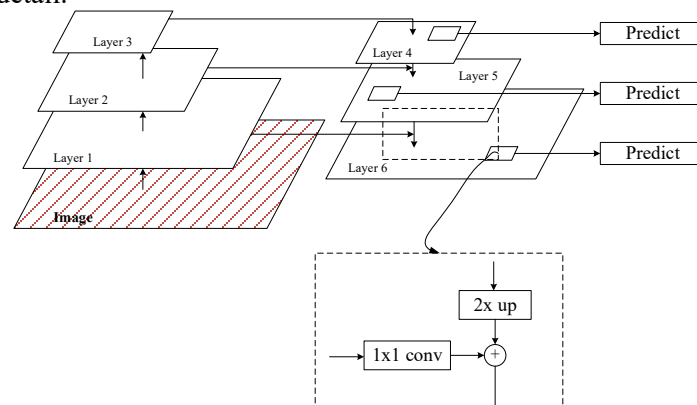


Figure 4. The work flow diagram of FPN

3.3. Number of scale prediction frameworks

The YOLOv3 algorithm uses the K-means[11] to cluster the target frame, and its objective function are equation (3) and equation (4).

$$J(\mu_1, \mu_2, \dots, \mu_k) = \frac{1}{N} \sum_{m=1}^K \sum_{n=1}^{N_m} (x_n - \mu_m) \quad (3)$$

$$\mu_m = \frac{1}{N} \sum_{n=1}^{N_m} x_n \quad (4)$$

where $\mu_1, \mu_2, \dots, \mu_k$ represent k clusters; N is the total sample size; N_m is the number of samples in the j th cluster center and m represents integers from 1 to K ; x_n represents an arbitrary sample and n represents integers from 1 to N_m .

We can learn from 1.1 that each prediction framework corresponds to an anchor box, and there are three cluster centers in each anchor box. In this paper, by selecting different K values and comparing the evaluation values of the algorithm (the specific method has been introduced in 3.1), the specific number of the best scale prediction framework for the optimizational YOLOv3 is determined. Take $K=6, 9, 12, 15$ respectively, and the change curve of evaluation value relative to iteration steps is shown in figure.5. When $K=6$, the evaluation value is far lower than other values; when $K=12$ is compared with $K=9$, the evaluation value is significantly improved, and $K=12$ is almost flush with the evaluation value of $K=15$. Therefore, with the increase of K , the evaluation value of the algorithm shows an upward trend, while with the increase of K , the evaluation value tends to be stable. When $K=15$, the training time of the model is much longer than $K=12$, which is not conducive to the engineering application. Therefore, this paper chooses $K=12$ to expand the original three scale prediction frameworks into four scale prediction frameworks, with the sizes of $13 \times 13, 26 \times 26, 52 \times 52, 104 \times 104$, respectively. Among them, 12 anchor boxes are $(10 \times 13), (16 \times 30), (33 \times 23), (32 \times 30), (40 \times 41), (30 \times 61), (62 \times 45), (93 \times 41), (59 \times 119), (116 \times 90), (156 \times 198)$ and (373×326) , respectively. Finally, the modified framework of YOLOv3 is shown in figure.7.

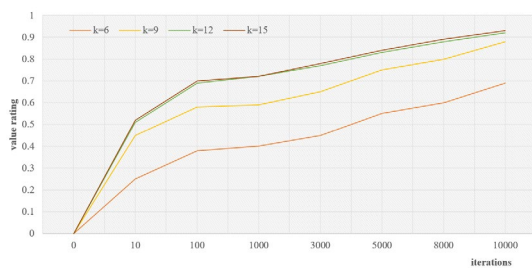


Figure 5. The change curve of evaluation value relative to iteration steps with $K=6, 9, 12, 15$

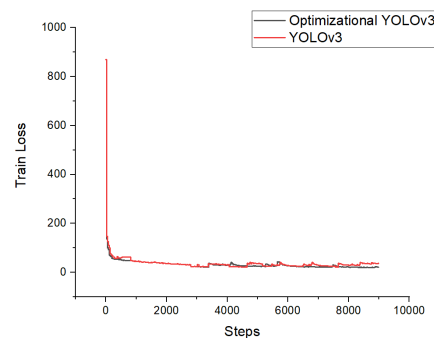


Figure 6. The scatter diagram of training loss values of original YOLOv3 and optimizational YOLOv3

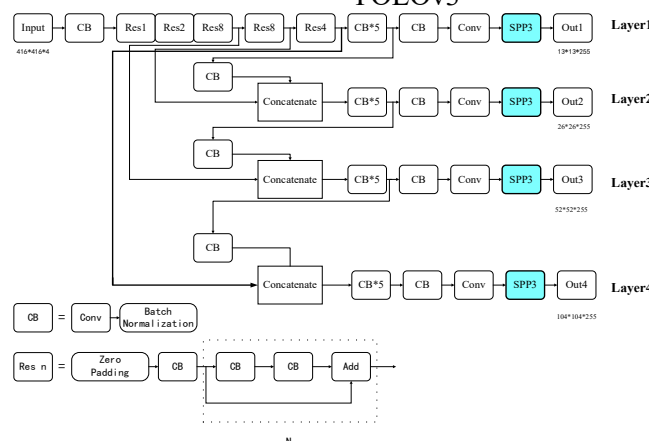


Figure 7. The modified framework of optimizational YOLOv3

4. Experiment results and analysis

4.1. Evaluation methods

In this paper, the results of weight calculation are taken as the evaluation value by using the three indexes of the missed detection rate, the false detection rate and the AP value. The specific algorithm is as follows.

$$\text{Evaluation_value} = (1 - \text{missed_detection_rate}) \times 30\% + (1 - \text{false_detection_rate}) \times 20\% + AP \times 50\% \quad (5)$$

$$X = \frac{T - M1}{T} \times 100\% \quad (6)$$

$$Y = \frac{M2 - M1}{M2} \times 100\% \quad (7)$$

where X is the missed detection rate; Y is the missed detection rate; AP is the average accuracy; T is the total number of devices marked as target class in the picture; $M1$ represents the total number of correct target equipment actually detected by the manufacturer; $M2$ represents the total number of target equipment actually detected by the manufacturer.

$$P = \frac{TP}{TP + FP} \quad (8)$$

$$R = \frac{TP}{TP + FN} \quad (9)$$

$$AP = \frac{1}{11} \sum_{R \in \{0.0, 0.1, \dots, 1\}} P(R) \quad (10)$$

where IoU takes 0.5 in process of calculating AP ; TP is the number of anchor boxes with $\text{IoU} > 0.5$; FP is the number of anchor boxes with $\text{IoU} \leq 0.5$; FN is the number of ground truth which is not detected.

4.2. Dataset selection and annotation

The dataset selected in this paper is divided into four categories: electric power fittings, foundations, towers and insulators. There are 50 types of defects, of which 32 are small size and not easy to find, such as suspension clamps rust, U-bolts rust, dampers rust, yoke plates rust, anchor clamps rust, U-bolts distortion, etc. We select 1247 UAV pictures to be taken as training sets, with different picture sizes in this paper. Software *Labelimg* is used to label the dataset manually, with a special number for each defect category, such as "hull of suspension clamp rust" labeled as "050101".

4.3. Training model and test results

The hardware environment for the algorithm in this paper is 16GB memory, NVIDIA GeForce RTX 2080 GPU, AMD Ryzen 7 2700X Eight-Core Processor and Ubuntu 16.04 operating system. 1247 training samples were trained with the Intersection-over-Union (IoU) of 0.5, batch size of 48, maximum iteration steps of 10,000, attenuation factor of 0.0005 and image resolution of 416*416. Training loss values are recorded during the training. At the same time, the same parameters are set to train the original YOLOv3 network, and the training loss values are recorded. The scatterplot of training loss values is drawn with the optimizational algorithm, as shown in figure.6. The figure shows that the training loss value of optimizational YOLOv3 is less than that of original YOLOv3 when the number of iteration steps approaches 10000, which proves that the improvement is effective.

We test the performance of the original YOLOv3 and the optimizational YOLOv3 with the same test set, and the test result of comparison used the test set with 489 pictures are shown in table 1. The contrast detection result of two algorithms run for several typical smaller sizes and the defects that is difficult to detect are shown in table 2 and figure.8. It is apparent that there are more faults have been detected in figure.8(b) and (d) than (a) and (c).

Table 1. The test result of original and optimizational YOLOv3 used the test set with 489 pictures

	Electric power fitting	Foundation	insulator	tower
YOLOv3 (evaluation value)	0.851	0.915	0.923	0.894
Optimizational YOLOv3 (evaluation value)	0.896	0.92	0.943	0.965

Table 2. The test result of original and optimizational YOLOv3 applied to electrical components

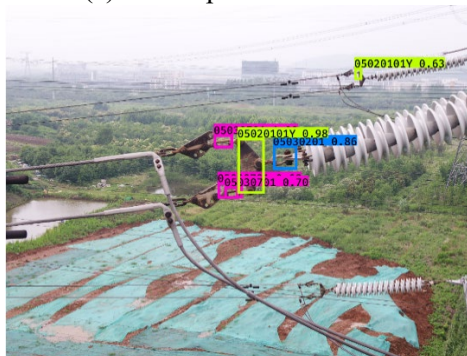
	Corrosive suspension clamps	Corrosive shackles	Loose aluminum armor tapes	Loose U-bolts	Tortile Grading rings
YOLOv3 (evaluation value)	0.765	0.795	0.652	0.354	0.245
Optimizational YOLOv3 (evaluation value)	0.863	0.864	0.756	0.512	0.426



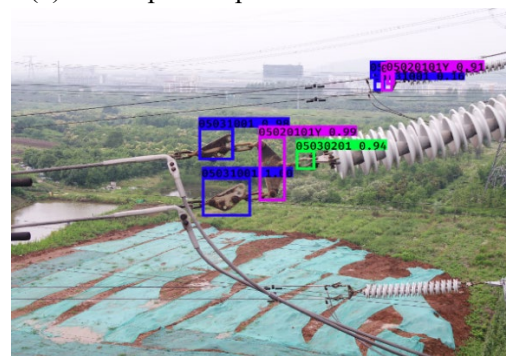
(a)The output of YOLOv3



(b)The output of optimizational YOLOv3



(c)The output of YOLOv3



(d)The output of optimizational YOLOv3

Figure 8. The contrast result of two algorithms run for several typical smaller sizes and the defects that is difficult to detect

According to table 1, the average evaluation value of the optimizational YOLOv3 on the test set including electric power fittings, foundations, towers and insulators has increased. Compared with the original YOLOv3 algorithm, the evaluation value on above types of defects has increased by 5.28%, 0.55%, 2.17% and 7.94% respectively. Among them, there are significant improvement in evaluation value of many smaller sized defects of power fittings and towers which cannot find easily. According to table 2, the evaluation values on hard to find and high-risk faults of four categories, such as critical

suspension clamps, critical shackles, loose U-bolts and tortile grading rings, are increased by 12.81%, 8.68%, 44.64% and 73.88% respectively in the power fittings category. And in the towers category, the fault evaluation value of loose aluminum armor tapes is increased by 15.95%. The experimental results show that the optimizational YOLOv3 algorithm not only improves the recognition rate of all faults, but also improves the recognition rate of small size and hard to find faults.

5. Conclusion

In this paper, an optimizational YOLOv3 based on SPP net and improved multi-scale framework is proposed. This algorithm can be applicable to the situation where exists the high rate of miss and false detection for the small and hard-to-find fault defects in the patrol inspection process of transmission system. It can improve such fault detection rate and the working efficiency of UAV in the inspection process. Although, this algorithm still has the problems such as high hardware requirements and long training time for the vehicle server. And the recognition rate for sundries near the foundations is still relatively not high. In the future, we will work to improve the accuracy of the algorithm in such aspects, and at the same time, without reducing the accuracy, reduce the hardware requirements of the algorithm on the vehicle server.

6. Reference

- [1] Wang B, Li W J and Tang H 2020 Improved YOLOv3 algorithm and its application in helmet detection *Computer Engineering and Applications* 11
- [2] Zhao Q, Li B Q and Li T W 2019 Target Detection Algorithm Based on Improved YOLOv3 *Laser & Optoelectronics Progress* 16
- [3] Deng Y N, Luo J X, Zhang Q, Liu Z, Hu Q, Jin F L and Bi P C 2019 Improved YOLOv3 network to evaluate air quality in images *Computer Engineering and Applications* 1
- [4] Lin G, Wang B, Peng H, Wang X Y, Chen S Y and Zhang L M 2019 *Electric Power Automation Equipment* **39** 213
- [5] Shashidhara S M and Sangameswara Raju P 2014 *IJEETC* **3** 34
- [6] Sondele S and Saini I 2013 *IJEETC* **2** 44
- [7] Ma P and Fan Y F 2020 Small Sample Smart Substation Power Equipment Component Detection Based on Deep Transfer Learning *Power System Technology* 12
- [8] Zhang F K, Yang F and Li C 2019 A fast vehicle detection method based on improved YOLOv3 *Computer Engineering and Applications* **55** 12
- [9] He K, Zhang X and Ren S 2016 *IEEE Conference on Computer Vision and Pattern Recognition* 770
- [10] He K M, Zhang X Y, Ren S Q and Sun J 2015 *IEEE transactions on pattern analysis and machine intelligence* **37** 1904
- [11] Wang C Y, Fan S S, Liu Z, Li B and Zhang W 2019 *Journal of Electric Power* **34** 322