A Hybrid Method Based on Object Detection and Image Augmentation for Substation Condition Monitoring

Junjie Chen[a,*], Siqi Pan[b], Yanping Chan[c], Yuedong Ni[d], Donghua Ye[e]

Fujian State Grid Zhangzhou Power Supply Company, Zhangzhou, 363000 Fujian, China

[a]755530740@qq.com, [b]2256345584@qq.com, [c]1527418424@qq.com, [d]283504278@qq.com, [e]18259610830@163.com

*Corresponding Author

Keywords: Substation condition monitoring, Yolov5, Multiscale data augmentation, Model pruning

Abstract: Artificial intelligence has made lots of achievements in the field of substation condition monitoring. However, in real world, due to the serious environment and small volume of data, the monitoring accuracy and the stability are poor. Therefore, a new method based on pruning YOLOv5 and spatial multiscale data augmentation (PYSMDA) is proposed for the substation condition monitoring. Firstly, an improved multiscale data augmentation method is proposed. According to different distribution, the spatial multiscale convolution is generated and used to enhance the defect image so that it increases the scale of data and the diversity of image data. The improved spatial multiscale data augmentation method weakens the interference of the varying environment on the recognition accuracy, and improves the defect detection accuracy. Then, the YOLOv5 is applied to train the multi-scale image data. For reducing the effect of the YOLOv5 large-scale parameters on the stability of the model, the model pruning method is utilized to shrink the structure parameters to improve the defect identification accuracy. The effectiveness of the proposed method is evaluated on the substation defect images. Experimental results indicate that the proposed method is well-performance for substation condition monitoring.

1. Introduction

Electric power is the key to ensure the normal operation of life, industry and other fields. The substation is the critical pole to realize power transmission in the power grid system. Due to serious working condition, there are many kinds of defects on substation equipment. Therefore, the manual condition monitoring method cannot adapt the rapid development of the power industry with long detection cycle and low detection accuracy[1]. With the development of artificial intelligence, new patrol inspection mode is developing more and more rapidly.

In recent years, several advanced monitoring devices are developed for substation condition monitoring, such as unmanned aerial vehicle (UAV)[2], infrared device[3], and so on. High-tech equipment promotes the development of remote analysis and defect recognition of data images, which needs the help of image recognition technology[4]. Ceron et al.[5] used Canny operator and direction adjustable filter to detect transmission lines. Khalayli et al.[6] extracted the edge features of insulators with gray matrix, classified the features with classifier, and detected the surface hydrophobicity of insulators. However, the above-mentioned methods need more advanced image processing technology to realize the extraction of defect features, and the applicability of the method is poor.

Image recognition based on deep learning technology has made a lot of breakthroughs. Deep learning can extract information with good performance from massive image data and automatically extract useful feature. Common deep learning methods include auto encoder (AE)[7], restricted Boltzmann machine (RBM)[8], convolutional neural network (CNN)[9], and so on. With the help of the characteristic advantages of CNN, the target detection method based on CNN is more and more popular[10]. Sermanet et al.[11] unified the classification and positioning tasks into a convolutional neural network through the Over Feat method, and used the multi-scale sliding window to detect each pixel position. Szegedy et al.[12] proposed the deep convolution neural network framework concept, and built a Googlenet model based on it. This model can increase the utilization rate of network parameters effectively.

With the continuous evolution of deep learning, object detection has made rapid progress. So many neural network models with good performance have emerged. The representative object detection algorithm includes two-stage network: R-CNN series[13], and one-stage network: Yolo series[14]. Dong et al.[15] proposed an improved lightweight YOLOv5 method for vehicle detection. They utilized C3Ghost and Ghost modules to reduce the floating-point operations (FLOPs). The experimental results indicated the progressiveness of the proposed method.

Although the deep learning method has good effects in the field of target detection, there are few studies in the field of substation condition monitoring. Therefore, it is promising to study the application of deep learning in substation condition monitoring. In the actual scene of substation defect detection, there are two main problems: (1) The shooting environment will change with the operation of the shooting tool, and this change of environment will affect the quality of the picture, thus affecting the accuracy of defect identification; (2) Because the field equipment usually does not have defects, the amount of data that can be captured in the defect fault image is very small, which leads to the model cannot complete the training completely, thus reducing the accuracy of defect recognition. Aiming at the above problems, a new method combining YOLOv5 pruning and spatial multiscale data enhancement is proposed to complete the substation equipment defect identification task. First, a spatial multiscale data enhancement method is proposed. This method uses different distributed initialization parameters to design convolution kernel, extract different spatial multiscale features from data. By applying spatial convolution, different distribution tricks are enhanced to improve the richness of data. Then, the model parameters are updated by YOLOv5 using the substation equipment defect images, and the model is shrunken by the model pruning method to obtain a simplified model.

The main contributions in this paper are summarized as follows.

(1) A novel method based on pruning YOLOv5 and spatial multiscale data augmentation is proposed for substation patrol inspection. In real application, the proposed detection method can get rid of harsh working environments and has adaptive capacity.

(2) A spatial multiscale data augmentation method is proposed for increasing data richness. A variety of different distribution methods are used to generate convolution kernel initialization weight parameters, and a variety of convolution kernels are generated by combining different-scale convolution kernels. In addition, spatial mechanism is adopted to rich the kinds of the convolution kernels. Therefore, the diversity of features is improved by extracting features through diversified convolution kernels.

(3) The pruning of the model is determined according to the weight distribution of the feature map, and the channels at the edge of the weight distribution are removed, so as to retain the channels with obvious feature extraction effect. The weight is used to measure the function of the channel, which can cut the channel more accurately and improve the efficiency of the model pruning.

The rest of this paper is arranged as follows. The related works about YOLOv5 applied in object detection in Section 2. In Section 3, the PYSMDA are described. The experimental results are displayed in figures and tables in Section 4. And in Section 5, the conclusions are drawn from experimental results.

2. Related Work

YOLO uses a single CNN model to complete the target detection task, which is divided into single-stage target detection algorithm. The updated version of YOLOv5 is adopted in this paper. It can be divided into four models, namely, YOLOv5s, YOLOv5m, YOLOv5l and YOLOv5x. The performance of these models is researched and obtained on the MS COCO test dataset as listed in Table 1. Due to the small amount of substation defect data, this paper uses YOLOv5s as the benchmark model. The structure of the YOLOv5s is shown in Fig. 1.

Table 1 the Comparison Performance of Different Models of Yolov5.

Model	Size(pixels)	mAP(0.5)	mAP(0.5:0.95)	Params(Mb)	FLOPs(G)
YOLOv5s	640 x 640	55.2	36.4	7.2	16.7
YOLOv5m	640 x 640	63.3	45.5	21.2	50.3
YOLOv5l	640 x 640	65.9	47.2	46.3	114.4
YOLOv5x	640 x 640	68.3	51.3	87.5	217.4

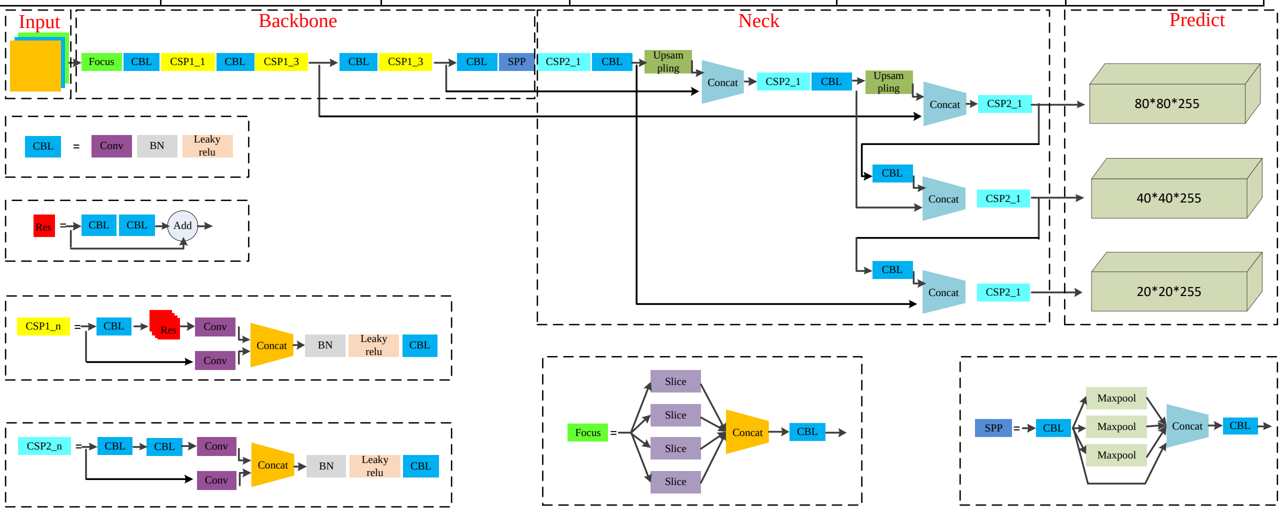


Fig.1 The Structure of the Yolov5s.

YOLOv5 consists of Input part, Backbone part, Neck network and Predict output part. Among them, the input end realizes mosaic data enhancement, adaptive anchor box calculation and picture scaling. Backbone network contains focus structure, convolution structure, C3 structure, spatial pyramid pooling structure and other modules. The input data is divided into four parts by focus module, and each part is equivalent to two down sampling. After the four parts of data are cut along the channel dimension, a binary down sampling feature map is obtained through convolution operation. The convolution structure includes the basic convolution unit that is consisting of convolution operation, activation function, and batch normalization. The neck network plays the role of generating pyramid features. By generating feature maps of different scales, the target detection effect is enhanced.

3. Proposed Method

3.1 Improved Spatial Multiscale Data Augmentation

Multiscale data augmentation is a classical data expansion method, which utilizes the commonality of the same type data. In general, multiscale data augmentation methods are adopted to mine more features from the raw data. According to the integrity of the raw data, multiscale features can be divided into local features and global features. Obviously, the global features are extracted from a whole visual angle of the data. And the local features are explored via a local visual field. With the help of CNN, the different size of the convolutional kernels can extract different scale features. The loss function is an accelerator for the convergence of the convolutional kernel. In the proposed method, however, the multiscale data augmentation method is only utilized as the data generator. Therefore, several different sizes of convolutional kernels are single on data generation due to the randomness of convolutional kernel initialization. In addition, spatial convolution is introduced to change the general convolutional structure, which helps to improve the feature learning capacity. A novel diversified spatial multiscale data augmentation is proposed to improve the richness of the generation data.

Four initialization rules are designed for spatial convolution, including random initialization, gaussian initialization, wavelet transform and plus rule. The details are shown in Fig. 2. The random initialization is to set the parameters of convolutional kernel in range of {0,1}.

Gaussian initialization uses gaussian kernel function to generate initialization parameters with the length of convolutional kernel. The formula is defined as follows:

$$Gaussian(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \qquad (1)$$

where μ is the mean of random variable x, and σ is the standard deviation of x.

The wavelet transform is adopted to initialize the parameters of the convolutional kernel. The initialization formula is defined as follows:

$$WT(a,\tau) = \frac{1}{\sqrt{a}}\int_{-\infty}^{\infty} f(t) * \varphi\left(\frac{t-\tau}{a}\right)dt \qquad (2)$$

where a is the scale and τ is the translation.

The plus rule applies plus operation to confuse above three initialization rules, and it is defined as follows:

$$plus = random + Gaussian + WT \qquad (3)$$

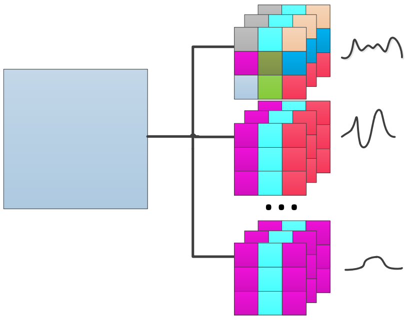


Fig.2 The Different Initialization Rules of the Improved Multiscale Data Augmentation Method.

3.2 Improved Pruning Yolov5

The big parameter storage and strong computing capabilities are indispensable for the implementation of the YOLOv5. However, the edge devices of the industrial application are usually unable to meet the demand. In addition, data scale of the industrial scene is smaller than life scene, and the current mainstream neural network cannot perform well with small data. Therefore, it is useful to reduce the size of the model without excessive loss of identification accuracy. Common pruning ideas are mainly divided into uniform channel pruning, convolution channel pruning and automatic pruning. The latter two kinds of pruning are designed for complex network structure and difficult detection tasks. Therefore, in this paper, the uniform channel pruning is adopted. The aim of the uniform channel pruning is to trim the convolutional kernel of the fixed channel at a proportion. The proportion is usually set by experience without considering the importance of different feature maps. Therefore, in this paper, a new uniform channel pruning method based on weight distribution of feature maps is proposed, as shown in Fig. 3.

A normal distribution is considered to constrain the proportion. It is defined as follows:

$$Normal(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \qquad (4)$$

3.3 Measurement Indicators

In common, to estimate the performance of the different object detection methods, precision, recall, mAP0.5, mAP0.5:0.95, average detection processing time, model size, FLOPs and parameter amount are utilized as the indicators.

Precision is calculated as follows:

$$Precision = \frac{TP}{TP+FP} \qquad (5)$$

Recall is defined as follows:

$$Recall = \frac{TP}{TP+FN} \qquad (6)$$

where TP is the positive instances correctly identified, FP denotes negative examples classified as positive instances, and FN is the positive samples that are incorrectly identified.

$$AP = \int_0^1 P(R)d\,R \qquad (7)$$

$$mAP = \frac{\sum_{i=1}^{N} AP_i}{N} \quad (8)$$

where mAP0.5 refers to the average AP of all categories when intersection of union (IoU) is set to 0.5, and mAP0.5:0.95 refers to the average mAP at different IoU values. The range of IoU can be changed from 0.5 to 0.95 with a step size of 0.05.

Given the prediction box D, its coverage area is defined as area (D). The truth box is G, and its coverage area is area (G). Then the IoU is defined as follows:

$$IoU = \frac{area(D) \cap area(G)}{area(D) \cup area(G)} \quad (9)$$

where the IoU denotes the aliasing degree between the prediction frame and the truth frame, and is a metrics of the quality of the generated prediction frame.

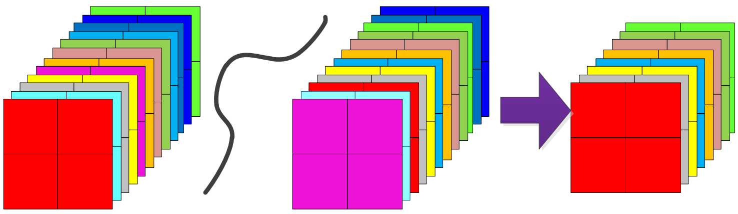


Fig.3 The Proposed Pruning Operation of Yolov5s.

4. Results and Discussions

4.1 Experimental Setting

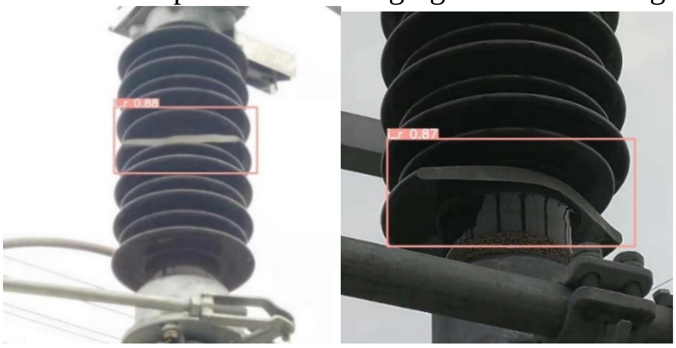
The experimental data includes the data collected from a substation and the data crawled by the network. The dataset includes five substation defects: fuzzy dial (f_d), damaged silica gel (dm_sg), insulator rupture (i_r), bird nest (b_n) and discolored silica gel (d_sg). The constructed substation condition monitoring dataset contains five categories of defect with a total of 7440 images. To train the model for substation condition monitoring, the dataset is divided into a training set, a validation set, and a testing set at a proportion of 8:1:1. Then, 5952 images are randomly selected as the training set, and the validation set and testing set are consisting of 744 images respectively. The dataset is shown in Fig. 4.

During the training phase, the Adam algorithm is utilized. For consistency, the learning rate of the optimizer is set to 0.001, and the epochs is set to 200.

Fig.4 The Dataset Obtained from Substation.

4.2 Result Analysis

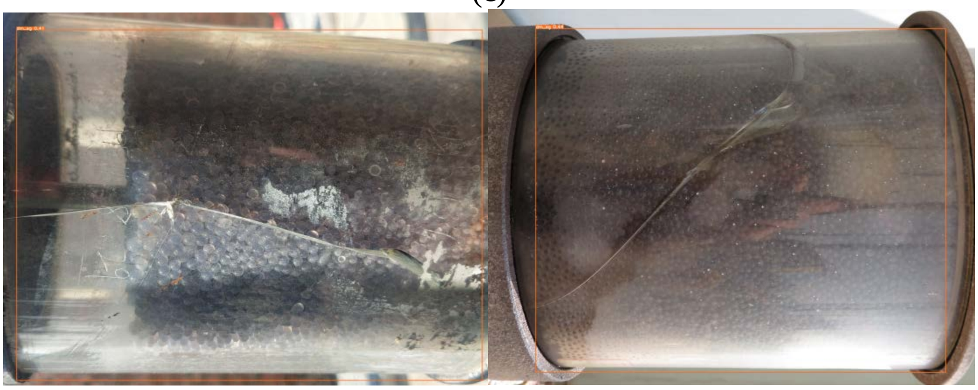
The detection results of different defects are shown in Fig. 5. Compared Fig. 5(a-e), we can see that the detection result of the damaged silica gel is the worst of the five defects. For showing the result visually, the single-type AP of five defects obtained from the PYSMDA are displayed in Fig. 6. From the result, we can see that the single-type AP of the discolored silica gel is the highest of the five defects with 98.54% and the detection result of the damaged silica gel is the worst with 40.32%. The main reason is that the damaged silica gel presents different texture structures, which are usually confused by obvious color. In addition, the damaged silica gel are often made of transparent glass, so it is difficult to distinguish crackles from other objects. As for the discolored silica gels, they are the same in the shape. And the changing colors of silica gels are distinguishable.



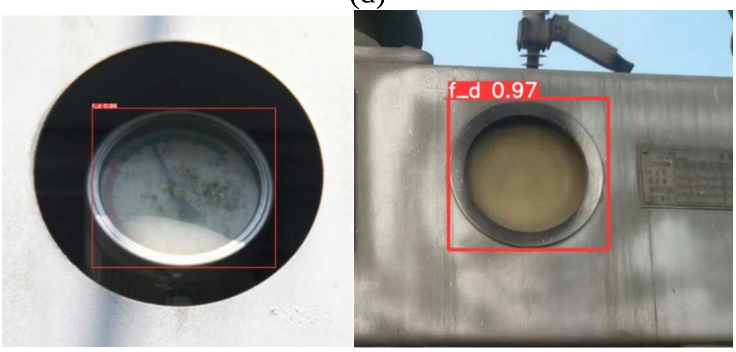
(a)

(b)

(c)



(d)



(e)

Fig.5 The Detection Result of Different Defects. (a) Insulator Rupture (B) Bird Nest (C) Discolored Silica Gel (d) Damaged Silica Gel (e) Fuzzy Dial.

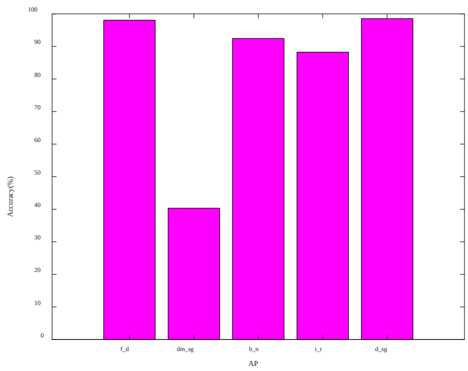


Fig.6 The Single-Type Ap of Different Defects.

Similarly, the single-type AP of the fuzzy dial is higher than the damaged dial. Because the fuzzy dial is only the color change in nature, but the damaged dial is changing in the structure. The changing structure increases identification difficulty and causes unsatisfactory results.

4.3 Ablation Experiment

In order to display the improved effect based on YOLOv5s, ablation experiments are implemented with four methods: YOLOv5s, pruning YOLOv5s, multiscale data augmentation + YOLOv5s and the proposed method.

The implementation details of other methods are shown as follows.

(1)YOLOv5s: In this method, images are input into the YOLOv5s without preprocessing stage. The learning rate is set to 0.001.

(2)pruning YOLOv5s: The proposed model pruning method is used to modify the YOLOv5s. The learning rate is set to 0.001.

(3) multiscale data augmentation + YOLOv5s: Compared with the proposed method, this method lacks pruning and improved multiscale data augmentation modules. The learning rate is set to 0.001.

The comparison results are shown in Table 2 and Fig.7. The results directly indicates that (1) The proposed pruning method can reduce the number of parameters and accelerate convergence of model with only a little loss at the accuracy. (2) The improved multiscale data augmentation shows nice performance, and the results proof the strong feature learning capacity of the improved multiscale data augmentation.

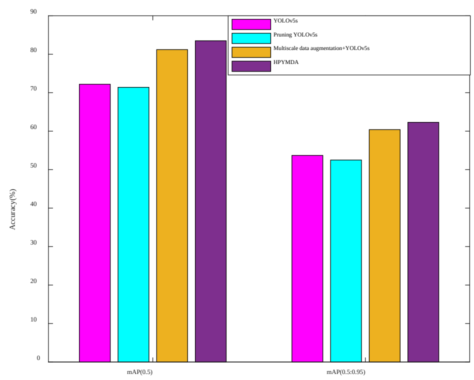


Fig.7 The Comparison Results of Different Methods.

Table 2 the Comparison Measurement Indictors of Different Methods.

Method	mAP(0.5)	mAP(0.5:0.95)	Params(Mb)	FLOPs(G)
YOLOv5s	72.2	53.7	7.2	24.3
Pruning YOLOv5s	71.4	52.5	6.4	22.5
Multiscale Data Augmentation + YOLOv5s	81.2	60.4	7.4	23.3
PYSMDA (ours)	83.5	62.3	6.7	22.4

Compare YOLOv5s and pruning YOLOv5s, the proposed model pruning method reduces model size with the difference of 0.8Mb. Moreover, the mAP(0.5) only reduces 0.7%. Therefore, the pruning method is effective for model pruning. With the improved multiscale data augmentation method, the mAP(0.5) increases by about 10%. And the improved technique is more useful than common multiscale data augmentation method. The powerful abilities of feature learning of the improved multiscale data augmentation are verified on the substation defect detection.

5. Conculsion

In this paper, a new method named PYSMDA based on pruning YOLOv5 and spatial multiscale data augmentation is proposed for substation condition monitoring. Different data distribution rules are applied for generating diversified weights as the initialization of the spatial convolution. Combined different initialization with different scale kernel size, the feature maps extracted via convolutional kernel are so abundant that improving the identification accuracy of the proposed method. A model pruning method based on weight proportion of different feature maps is proposed for cutting parameters of the YOLOv5. Compared with simple and rude model pruning methods by reducing number of channels, the proposed pruning method implements parameters reduction reasonably. The effectiveness of the proposed method is verified on dataset collected from substation, and the results indicates that the PYSMDA is useful and valuable for substation condition monitoring in real application.

Acknowledgement

The authors acknowledge the 2022 science and technology project of State Grid Fujian Electric Power Co., Ltd, Research on substation UAV image technology based on deep learning B3135022000C.

References

[1] Wang, Z. W. et al. Combined inspection strategy of bionic substation inspection robot based on improved Biological Inspired Neural Network, Energy Reports. 7 (2021) 549-558.

[2] Lekidis, A., Anastasiadis, A. G., Vokas, G. A. Electricity infrastructure inspection using AI and edge platform-based UAVs, Energy Reports. 8 (2022) 1394-1411.

[3] Gadd, H., Werner, S. Fault detection in district heating substations, Applied Energy. 157 (2015) 51-59.

[4] Ahmad, H. M., Rahimi, A. Deep learning methods for object detection in smart manufacturing: A survey, Journal of Manufacturing Systems. 64 (2022) 181-196.

[5] Ceron, A., Mondragon, B. I. F., Prieto, F. Power line detection using a circle based search with UAV images, 2014 International Conference on Unmanned Aircraft Systems (ICUAS), 2014, pp. 632-639.

[6] Khalayli, L. et al. Automatic inspection of outdoor insulators using image processing and intelligent techniques, 2013 IEEE Electrical Insulation Conference (EIC). 2013, pp. 206-209.

[7] Wang, Y. R., Sun, G. D., Jin, Q. Imbalanced sample fault diagnosis of rotating machinery using conditional variational auto-encoder generative adversarial network, Applied Soft Computing. 92 (2020), 106333.

[8] Chopra, P. K., Yadav, S. Restricted boltzmann machine and softmax regression for fault detection and classification, Complex & Intelligent Systems, 4 (2018), pp. 67-77

[9] Liang, P., Deng, C., Wu, J., Yang, Z. Intelligent fault diagnosis of rotating machinery via wavelet transform, generative adversarial nets and convolutional neural network, Measurement 159 (2020) 107768.

[10] Qi J. T., Liu, X. N., Liu, K. An improved YOLOv5 model based on visual attention mechanism: Application to recognition of tomato virus disease, Computers and Electronics in Agriculture, 194 (2022) 106780.

[11] Sermanet, P. et al. OverFeat: Integrated Recognition, Localization and Detection using Convolutional Networks, arXiv preprint arXiv:1312.6229, 2013.

[12] Szegedy, C. et al. Going Deeper with Convolutions, 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015, pp. 1-9.

[13] Wang, D. D., He, D. J. Fusion of mask RCNN and attention mechanism for instance segmentation of apples under complex background, Computers and Electronics in Agriculture, 196 (2022) 106864.

[14] Liu, T., Li, G. L., Gao, Y. Fault diagnosis method of substation equipment based on You Only Look Once algorithm and infrared imaging, Energy Reports 8 (2022) 171-180.

[15] Dong, X. D., Yan, S., Duan, C. Q. A lightweight vehicles detection network model based on YOLOv5, Engineering Applications of Artificial Intelligence, 113 (2022) 104914.