2024 Asia-Pacific Conference on Software Engineering, Social Network Analysis and Intelligent Computing (SSAIC) | 979-8-3503-9368-2/24/$31.00 ©2024 IEEE | DOI: 10.1109/SSAIC61213.2024.00009

Image Recognition and Processing Algorithm of Power Grid Facilities Map Based on Deep Learning

Li Wang[a], Fei Chen[b], Yuxiang Wang[c], Ying Liu[d], Chen Luo[e]
Guizhou Power Grid Company Limited, Guiyang, Guizhou, 550002 China
[a]37331309@qq.com, [b]81772141@qq.com, [c]520010548@qq.com, [d]64657584@qq.com, [e]lchen2005@qq.com

***Abstract*—The traditional image processing method of power grid facilities map is based on iconography, which can alleviate the artificial pressure to a certain extent. However, due to the slow speed and low accuracy of the traditional iconography method, it is difficult to be applied in the field of fault inspection. In order to realize intelligent power inspection more quickly and accurately, an image recognition and processing algorithm of power grid facilities map based on deep learning is proposed to solve the problems of occlusion, inaccurate classification and insufficient feature extraction in the actually collected power grid facilities map images. The convolution operation module and residual module in YOLOv5 algorithm are improved, and the learning depth of the algorithm is deepened by increasing the number of convolution layers. At the same time, the SENet attention mechanism is added to the basic convolution module. The research results show that the accuracy of this model for power equipment identification has reached more than 99%. And the recognition accuracy of fault defects can reach 92.746%. This model improves the detection accuracy and speed of power grid facilities map images, and also provides a novel and feasible scheme for intelligent detection of power grid facilities map images.**

Keywords—Power Grid, Facilities Map, Image Recognition, Deep Learning, YOLOv5

I. INTRODUCTION

With the rapid development of AI (artificial intelligence) technology, more and more industries introduce AI technology, and the power industry is no exception[1]. And the application of image recognition is even more obvious. Image recognition technology can not be separated from the fingerprint card punching machine used by mobile phones, computers, etc. to the sign-in system of face recognition. As one of the important branches of AI technology, target detection combines target location with target classification, and uses multi-directional knowledge such as image processing technology and machine learning to locate interested objects from images or videos[2]. At present, target detection has been widely used in many fields, such as automatic driving, unmanned aerial vehicle inspection, product quality inspection, video surveillance and so on[3-4].

The traditional image processing method of power grid facilities map is based on iconography, which can reduce the artificial pressure to a certain extent, but it is difficult to be applied in the field of fault inspection because of the slow speed and low accuracy of the traditional iconography method[5-6]. Literature[7] puts forward two algorithms for insulator identification of gray-scale images. The first is the segmentation algorithm of local feature matching combined with active contour model, and the second is the insulator detection algorithm based on texture features, local features and spatial distribution features. Literature[8] Based on the Faster RCNN algorithm, seven kinds of substation equipment, such as transformers, bushings and circuit breakers, are detected to achieve accurate positioning and identification of the equipment. Literature[9] improved the traditional region growing segmentation method, extracted Hu invariant moments of substation equipment as feature vectors, and then used support vector machine to identify infrared images of substation equipment.

Under the background of building smart grid and enhancing the reliability of power supply, it is the general trend to replace manual inspection with unmanned aerial vehicles. Traditional inspection image analysis techniques are mostly based on classical image processing algorithms, and the manual features extracted by these algorithms basically belong to the underlying visual features. Compared with the features of pure learning, the features are more interpretable but less adaptable to data[10]. With the wide application of deep learning technology in the field of computer vision, object detection based on deep learning has reached an unprecedented height. As a representative of single-stage detection algorithm, YOLOv5 algorithm has the advantages of less code, simple program, high detection speed and high detection accuracy. Based on YOLOv5 structure, this paper improves the convolution operation and residual operation in the original algorithm to solve the problems of long detection time and excessive consumption of computer hardware resources in the current mainstream power grid facility map image recognition and processing algorithms.

II. RESEARCH METHOD

A. YOLOv5 algorithm principle

Deep learning stands at the crossroads of multiple disciplines, including neural networks, artificial intelligence (AI), graphical modeling, optimization, pattern recognition, and signal processing. Drawing inspiration from the human brain's intricate and hierarchical structure, deep learning endeavors to emulate the cranial nerve system. It achieves this by establishing a layered model framework that resonates with the way our brain functions. This structure allows for a sequential extraction of input data, gradually morphing it into a more abstract high-level representation, whether that's a specific attribute, category, or feature. The profound nature of deep learning can be better understood when diving into its high-level descriptive attributes. Two pivotal aspects emerge that define its core:

Hierarchical Processing: The framework is designed in a manner that information undergoes multi-tiered or staged nonlinear processing. This means that as data progresses

979-8-3503-9368-2/24/$31.00 ©2024 IEEE
DOI 10.1109/SSAIC61213.2024.00009

through each layer, its transformation is non-uniform, making it adaptable to a variety of applications. The structure facilitates the capture of intricate patterns or anomalies that a linear approach might miss.

Evolving Feature Representation: As deep learning models delve deeper into layers, the methodology for learning these features, whether supervised or unsupervised, becomes increasingly abstract. In essence, the complexity and abstraction of feature representation amplify as one progresses from one layer to the next. This multilayered approach ensures that the system extracts increasingly nuanced and detailed insights from the data.

In sum, deep learning offers an advanced computational model that mirrors human cognitive processes, enabling machines to dissect, interpret, and act upon data with remarkable precision and depth.

At present, YOLOv5 is widely used in industry because of its advantages such as fast, accurate and light weight. YOLOv5 and YOLOv4 are basically similar in structure, but slightly different in details[11]. Compared with YOLOv4, YOLOv5 network adds a Focus structure, and the image is sliced through the Focus structure, and the double down-sampled feature map is obtained without information loss, which effectively reduces the feature loss. YOLOv5 model structure is shown in Figure 1.

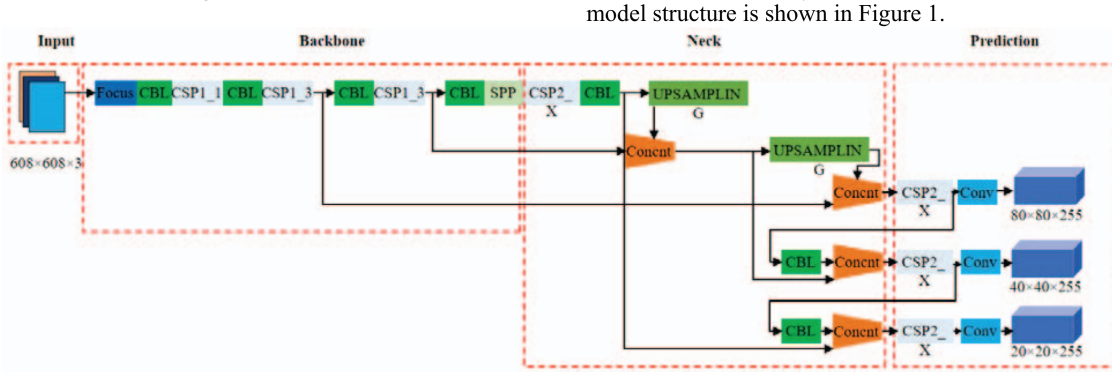

Figure 1 YOLOv5 model structure

The main principle of Focus module is slicing operation, which is similar to the double down-sampling operation with convolution kernel of 2 * 2 and step length of 2. Based on the previous version, YOLOv5 further improves the recognition accuracy of the algorithm by using the new LOSS function and bottleneck structure.

At present, YOLOv5, as a widely used algorithm model, achieves a balance between recognition accuracy and recognition speed, and is more suitable for defect image target detection in the scene of power grid inspection compared with other networks.

B. Image recognition and processing of power grid facilities map

Because the training of image detection model of power grid facilities map needs the same format data set, in order to facilitate data processing, it is necessary to normalize the image first. Accord to that normalization processing and los function processing, the large errors generated in conversion and proces are reduced, so that the image features are preserved as completely as possible[12-13]. The purpose of image normalization is to unify the format and size of the image without changing the pixel value.

When sample data is stationary, meaning each data dimension's statistics adhere to a consistent distribution, normalization can be achieved by deducting the data's statistical average from every sample. In the context of images, such a normalization method can nullify the average brightness or luminosity, as often, the image content holds greater significance than its ambient light levels. The specific procedure for feature standardization involves initially determining the mean value for data across all dimensions. Subsequently, this mean value is deducted from each dimension, followed by dividing each dimension's data using its respective standard deviation. To ensure each feature component's influence is balanced, standardization typically applies individually to every component.

If one were to directly employ a model for training, the detection outcome might be compromised due to discrepancies in detection target scales and sample counts. A viable strategy is categorizing training images into multiple groups, individually training them using various models, and then deducing results uniformly. By excluding categories with limited quantities and subpar label quality, several smaller and indistinguishable categories can be consolidated.

Within YOLOv5, apart from fundamental data enhancement techniques, the Mosaic data enhancement approach is also utilized. Its core principle is straightforward: it entails the random cropping, scaling, and subsequent rearrangement of four images to produce a single composite image. This method augments dataset objectives, amplifies smaller target sample counts, and expedites network training. The refined YOLOv5 algorithm is employed for detecting and pinpointing image targets within power grid facility maps. The configuration of this enhanced algorithm is depicted in Figure 2.

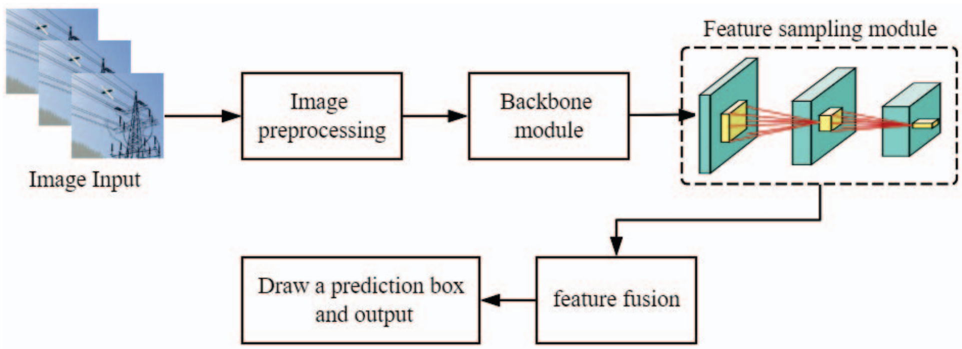

Figure 2 Improved YOLOv5 algorithm structure

Human attention reflects a distinct cerebral signal processing attribute intrinsic to our visual perception. By swiftly perusing an entire image, humans identify areas warranting heightened focus. Once these attention-centric zones are discerned, greater attention is directed to them, enabling acquisition of intricate target details while concurrently filtering out extraneous information. The attention mechanism closely mirrors this human visual focus process. In this study, the SENet module is explored, with Figure 3 illustrating the operational essence of the SE module.

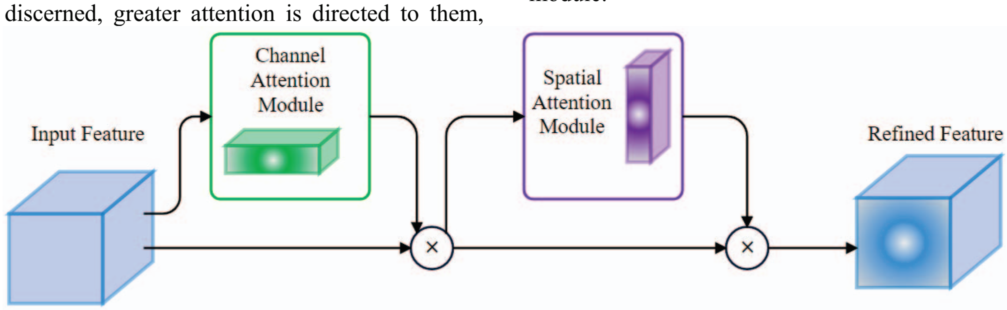

Figure 3 SENet module

SENet module mainly includes Squeeze and Excitation. The aggregation is followed by the excitation operation, which adopts a simple self-gating mechanism, which takes embedding as input and generates a set of corresponding weights for each channel.

F_{sq} structure compresses the input feature matrix U in space dimension, that is, global average pooling operation, and the formula is shown in (1).

$$z_c = F_{sq}\left(u_c\right) = \frac{1}{H*W}\sum_{i=1}^{H}\sum_{j=1}^{W} u_c\left(i,j\right) \quad (1)$$

Global average pooling compresses $H*W$ eigenvalues on each channel into an average, which is the representative of all eigenvalues of the channel and has a global receptive field. Therefore, the input characteristic matrix U becomes the characteristic matrix of $1*1*C$ after the compression structure, and is input into the following excitation mechanism.

In this paper, SENet attention mechanism is introduced into the bankbone of YOLOv5. In the feature extractor Goog Le Net, the Inception module embeds the SENet attention mechanism framework. In addition to Goog Le Net feature extractor, SENet attention mechanism framework can also be embedded in other classical feature extractors, such as Res Net and Dark Net [14].

YOLOv5 replaces the focus layer of the first layer of the network with the convolution layer, but it is found through experiments that it will lead to the decline of the detection accuracy of small targets. Divide the image into blocks and process them one by one. This operation can lossless down-sample, effectively reduce the feature loss of small targets, and effectively improve the detection accuracy of small targets. Therefore, the convolution of the first layer of the latest version of YOLOv5 is still replaced by the old version of Focus layer to improve the detection accuracy of small targets.

Feature fusion usually treats all the different feature maps input into the network equally, but because the feature maps of different scales have different resolutions, their contributions to the output fusion features are different. In order to solve this problem, the weighted two-way feature pyramid adds an extra weight to each input and lets the network learn the importance of each input. Therefore, fast normalized fusion is introduced into BI-FPN, and its formula is as follows:

$$O = \sum\nolimits_I \frac{w_i}{e+\sum\nolimits_j w_j} * I_i \quad (2)$$

Among them, $w_i \geq 0$ sets the learning rate $e_i = 0.0001$ to a small value by applying the RELU activation function after each w_i to avoid numerical instability. Similarly, the value of each normalized weight is also between [0,1], but because there is no Softmax operation here, the efficiency is much higher.

The residual network is used to predict the position of the target in the feature map to construct the feature pyramid. Fig. 4 shows the improved sub-module. In insulator identification, the image background is complex, and it is difficult to distinguish effectively by background color difference, so the residual edge is also convolved. CSMA operation is used to improve the accuracy of target feature extraction and alleviate the problem of image gradient disappearing due to the increase of learning depth.

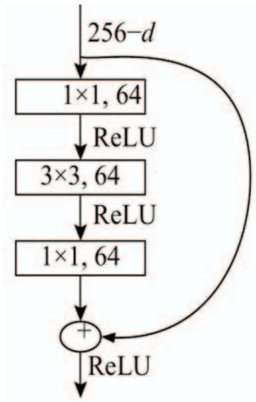

Figure 4 Improved residual module

The loss function of this paper adopts the improved $DIOU\_NMS$ as the loss function, which replaces the $GIOU$ loss function with slow convergence, and effectively solves the problem that the map image of power grid facilities to be detected is occluded. The loss calculation formula of $DIOU$ is:

$$L_{DIOU} = 1 - 0.5 + R_{DIOU}\left(B, B^{gt}\right) \quad (3)$$

Where: $R_{DIOU}\left(B, B^{gt}\right)$ is the penalty term of prediction frame B and target frame B^{gt}.

III. Experimental results and analysis

Taking the map image data set of power grid facilities of a company in China Southern Power Grid as an example, the experiment is carried out. When establishing the image sample set, the influence of different environmental factors is fully considered, and different representative types of images are selected for experimental analysis. Covering the types of missing pins with different angles and light to improve the accuracy of line defect detection. Randomly select 7000 of them as training set and 3000 as test set, and summarize all the information into CSV file. After the data labeling is completed, the images are preprocessed, and the features of all images are extracted based on HIS color model to enhance the image quality.

The test computer is configured as CPU Intel Core i5—6600k, GPU is NVIDIA GTX1080Ti, memory is 16 G, and the system is ubuntu 16.04. The resolution of the input picture is 478*478, the batch_size is set to 8, the optimizer chooses Adam, and the cosine annealing attenuation method is used to adjust the learning rate. The initial and cut-off learning rates are 1e-4 and 1e-6 respectively, and the training epoch in each stage is 100.

In order to better evaluate the quality of the improved model, this paper introduces AP evaluation index. Fig. 5 mainly compares the detection results of this scheme, Faster RCNN, YOLOv3 and YOLOv4.

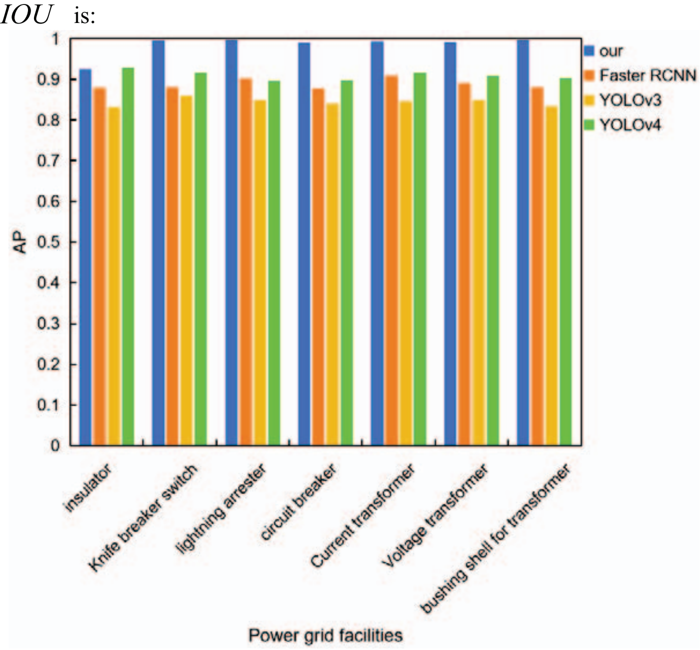

Figure 5 Accuracy of power equipment identification with different algorithms

Because most of the seven kinds of power equipment in this data set are small targets, it increases the difficulty of overall detection, and the accuracy of insulators is only 0.926%, which is prone to missed detection and false detection. The other six kinds of power equipment have reached more than 99%.

When the smoothing L1 loss function is used for location loss and the Logistic cross entropy loss function is used for category loss, the target detection is carried out on the map image of power grid facilities to judge the effectiveness of the proposed algorithm. Table I shows the comparison results of several target defect detection based on deep learning.

TABLE I. COMPARISON OF FAULT DEFECT IDENTIFICATION ACCURACY

Algorithm	F1	Processing time/s	Accuracy (%)
Faster RCNN	0.912	4.277	88.021
YOLOv3	0.815	5.48	84.722
YOLOv4	0.881	4.027	89.75
Proposed scheme	0.958	2.936	92.746

As can be seen from the table, with the optimization and improvement of the network model, the recognition accuracy of the picture is constantly improving. Compared with other similar methods, the improved deep learning algorithm proposed in this paper has achieved an ideal target detection effect on test data sets, with the shortest real-time processing time, and the recognition accuracy of fault defects can reach 92.746%.

IV. CONCLUSIONS

As a representative of single-stage detection algorithm, YOLOv5 algorithm has the advantages of less code, simple program, high detection speed and high detection accuracy. Based on YOLOv5 structure, this paper improves the convolution operation and residual operation in the original algorithm to solve the problems of long detection time and excessive consumption of computer hardware resources in the current mainstream power grid facility map image recognition and processing algorithms. The research results show that the accuracy of this model for power equipment identification has reached more than 99%. With the optimization and improvement of the network model, the recognition accuracy of pictures is constantly improving. Compared with other similar methods, the improved deep learning algorithm proposed in this paper has achieved an ideal target detection effect on test data sets, with the shortest real-time processing time, and the recognition accuracy of fault defects can reach 92.746%. However, the algorithm described in this paper uses multiple networks to extract features, solves the problem of recognition accuracy through feature fusion, and increases the redundancy of the model. In the next step, we should study a more simplified network feature extraction model to realize end-to-end training.

REFERENCES

[1] Liu, X. , Dong, B. , Li, P. , Yuan, B. , & Wang, K. (2022). Research and application of image recognition of substation inspection robots based on edge computing and incremental learning. Nonlinear Optics, Quantum Optics(1/2), 56.

[2] Chen, J. H. , Su, M. C. , Cao, R. , Hsu, S. C. , & Lu, J. C. (2016). A self organizing map optimization based image recognition and processing model for bridge crack inspection. Automation in Construction, 73(7), 58-66.

[3] Satoh, H. , & Takeda, F. (2012). System for image-processing-based inspection of a screw hole on a molded plastic frame. Ieej Transactions on Industry Applications, 132(7), 701-708.

[4] Martinez, C. , Sampedro, C. , Chauhan, A. , Collumeau, J. F. , & Campoy, P. (2018). The power line inspection software (polis): a versatile system for automating power line inspection. Engineering Applications of Artificial Intelligence, 71(8), 293.

[5] Li, J. , Wang, Q. , & Li, M. (2017). Electric equipment image recognition based on deep learning and random forest. Gaodianya Jishu/High Voltage Engineering, 43(11), 3705-3711.

[6] Jin, L., Wang, H., Wang, W., & Yan, S. (2018). Transmission line fault location based on machine vision and laser ranging. Tongji Daxue Xuebao/Journal of Tongji University, 46(12), 1745-1753.

[7] Liu L, Du R, Liu W. Flood Distance Algorithms and Fault Hidden Danger Recognition for Transmission Line Towers Based on SAR Images. Remote Sensing. 2019; 11(14):1642. https://doi.org/10.3390/rs11141642.

[8] Hu, D. , Zhu, J. , Liu, J. , Wang, J. , & Zhang, X. (2022). Gesture recognition based on modified yolov5s. IET image processing(8), 16.

[9] Fang, Y. , Wang, Y. , Liu, C. , & Liu, X. (2022). Near-field radio frequency identification leaky-wave antenna based on artificial transmission line. International journal of RF and microwave computer-aided engineering(4), 32.

[10] Li, B. , & Chen, C. (2018). Transmission line detection based on a hierarchical and contextual model for aerial images. Journal of Electronic Imaging, 27(4), 1.

[11] Peng, X. , Liang, F. , Qian, J. , Yang, B. , & Zheng, X. (2019). Automatic recognition of insulator from uav infrared image based on periodic textural feature. Gaodianya Jishu/High Voltage Engineering, 45(3), 922-928.

[12] Lin, T. , & Liu, X. (2021). An intelligent recognition system for insulator string defects based on dimension correction and optimized faster r-cnn. Electrical engineering(1), 103.

[13] Li, Z. T. , Hao, Y. P. , Li, L. C. , Yang, L. , & Fu, C. (2011). Image recognition of ice thickness on transmission lines using remote system. Gaodianya Jishu/High Voltage Engineering, 37(9), 2288-2293.

[14] Wang, D. , & He, D. (2021). Channel pruned yolo v5s-based deep learning approach for rapid and accurate apple fruitlet detection before fruit thinning. Biosystems Engineering, 210(6), 271-281.