Research on the use of YOLOv3 image processing algorithm in power plants

Junjie Zhao*

Institute of computer and control engineering, North China Electric Power University, Baoding 071000, China

*Corresponding author e-mail: 1224073665@qq.com

Abstract. The power system is a guarantee system for nation-building, and it is inseparable from the use of electricity everywhere in life. There are many links in the production of power plants that need to be well monitored, but good real-time monitoring has always been a problem. The YOLOv3 algorithm draws on the residual network structure in the target detection of image processing, forms a deeper network level, and multi-scale detection, which improves the detection of mAP and small objects. Applying the YOLOv3 algorithm to the image processing of power plants can effectively improve the efficiency of image detection in power plants. Based on the different network structures of YOLOv3, this paper studies the power plant image processing in terms of detection speed and accuracy. The research shows that although the results of different network structures are different, the application of YOLOv3 in power plant image processing is still at the detection speed. Fast, and its detection accuracy is also very high.

1. Introduction
The power plant converts raw energy (such as water, steam, diesel, gas, etc.) into electricity to provide electricity for our lives. Among all sustainable power generation technologies, combustion power generation is still the main mode in China. Among them, coal-fired power plays a major role in combustion power generation. The coal-fired power generation process includes the powder supply, combustion, power generation and auxiliary links. The entire power generation process requires real-time process monitoring. It is very meaningful to control each link in the production process of the power plant. The changes in each link have an impact on production cost, production efficiency, and production safety. The traditional real-time monitoring of power plants is generally monitored by the Power Plant Decentralized Control System (DCS). However, in the DSC monitoring process of the power plant, the monitoring and control strategy is based on relevant extended data from indirect measurements, such as main air temperature, main gas pressure, etc. However, the monitoring method by indirect measurement control is not reliable. For example, the furnace flame is associated with all the process parameters, which are related to the safety, stability and reliability of all aspects of the power generation process and production. The real-time process monitoring of objects with complex factors such as the furnace flame cannot be controlled by traditional indirect data measurement, so the advantages of image processing technology are reflected in this aspect.
Image processing can be used in power plants to better control and visualize the production process of power plants. The application of digital image processing technology to boiler flame detection has opened up a new field for power plant boiler flame detection. The combination of digital image processing technology and computer technology for flame detection is a new method of flame detection. As an important part of the flame detection system, the flame image sensor is designed for the shortcomings of the previous flame image sensor. Xu Bing designed a flame image sensor using the image fiber and CCD camera, and compressed the image to save the image. Storage space and transmission time, which laid the foundation for remote monitoring of flame [1]. By measuring the flame condition of the flame image processing and combining the detection and control technology of the flame developed by the expert system, the efficiency and energy saving effect of the boiler combustion system can be improved. Ji Changan, Zhang Xiubin and others have realized the application research of image processing technology and fuzzy control in power plant reconstruction through the methods of detecting, processing, feature extraction and pattern recognition of large-scale boiler multi-angle burners. The combustion efficiency is obviously improved compared with the previous introduction control method, and the high-power electrical equipment for the process of air supply, air intake, coal feeding and the like is optimized on the basis of the combustion control, so that the process meets the boiler-to-coal ratio and the furnace negative pressure. Optimal control of electrical energy is obtained under controlled conditions such as superheated steam and pressure [2].

In the medium detection of power plant, the image processing technology can be used to automatically identify the car part of the coal car. On this basis, the sampling point is selected by the program to select the sampling point for sampling, which ensures that the sampling point selected by the head has been identified. Within the range of the carriage, the sampling work is guaranteed to proceed normally [3]. In the maintenance of power line inspection, image processing is used to check the appearance of insulators and attached objects. It is good to judge whether the insulators installed outdoors and the input image are subject to sunshine changes are intact [4]. Image processing can be said to play an important role in the production process of power plants. The ability of image processing (processing speed, accuracy, etc.) also affects the production of this power plant. Image processing based on YOLOv3 algorithm has great advantages in speed and precision compared to other image processing technologies.

YOLO algorithm English full name You Only Look Once, is an article of CVPR2016, well-known in the field of target detection. The YOLO algorithm can recognize all objects by looking at the picture once, and real-time object detection, which is about 40 frames per second, is very fast. After YOLO, YOLOv2 and YOLOv3 were improved. The accuracy after the improvement was greatly improved compared with v1. Many image processing areas now utilize the YOLOv3 algorithm. Liu Wei and Guo Lijun proposed a deep neural network based on the improved YOLOv3 algorithm to solve the problem of poor positioning accuracy and low efficiency of the traditional container box number, and realize the rapid positioning of the container box number. Based on the improved YOLOv3 algorithm, the container box number positioning method has high accuracy and strong real-time performance, and the positioning accuracy rate is 98.5%, and the positioning rate of 26.23 fps can be reached [5]. On the target tracking, the target tracking algorithm based on the deep learning detection algorithm YOLOv3 has good robustness in complex situations such as target illumination changes, attitude changes, dimensional changes, and rotational changes. Li Jing and Huang Shan studied the target tracking algorithm based on the YOLOv3 target tracking method. The research shows that the target tracking method based on YOLOv3 has a good tracking effect, and its comprehensive performance is improved by 80 compared with the four comparison algorithms in the experiment. % or so [6]. For the problem of free parking of shared bicycles, the multi-scale detection training and k-means dimension clustering are used to improve the YOLOv3 network to obtain the feature matrix of the shared bicycle, thereby calculating the overall running state of the shared bicycle and performing state statistics to obtain the bicycle. Location and status [7]. The multi-scale clustering convolutional neural network MK-YOLOv3 algorithm avoids the mutual occlusion between humans and objects, the inaccurate detection of small targets and the influence of complex illumination intensity on pedestrian detection, in order to realize the recognition and detection of pedestrians in VOC data. The accuracy and speed of small target recognition on the set
are greatly improved [8]. In the research of target recognition algorithm by Liu Xueping and Li Yugan, in order to accurately identify the target parts in the image, an improved YOLOv3 target recognition algorithm based on adaptive edge optimization is proposed. The improved algorithm has improved the precision and improved the overall performance of the network [9].

The YOLOv3 algorithm has great advantages in image processing. It can be seen in various types of image processing applications, but it has not been used in power plant applications. The use of YOLOv3 image processing algorithms in power plants is of great practical value. Changes in various factors in the production of power plants will affect production efficiency and production costs, and even serious production accidents will occur. If the YOLOv3 algorithm is applied to the image processing of a power plant, the processing effect will inevitably increase a lot.

This paper uses different network structures to study the use of YOLOv3 image processing algorithms in power plants in terms of speed and accuracy. The study shows that the network structure of different YOLOv3 algorithms has some differences in speed and accuracy in image processing of power plants, but its advantages are still significant compared with other image processing.

2. Methods

2.1. YOLOv3 algorithm

The basic idea of the YOLO algorithm is to extract features from the input image through the feature extraction network, and obtain a feature map of a certain size, such as the input image is 13*13, and then divide the input image into 13*13 grid cells, and then if ground truth If the center coordinate of an object falls in which grid cell, the grid cell predicts the object, and each grid cell predicts a fixed number of bounding boxes (2 in YOLO v1 and 5 in YOLO v2. There are three in YOLO v3. The initial size of these bounding boxes is different. The only bounding box of the IOU in the bounding box is the one used to predict the object. The predicted output feature map has two dimensions that are the dimensions of the extracted features, such as 13*13, and one dimension (depth) is B*(5+C). Note: YOLO v1 is (B*5+ C), where B represents the number of bounding box es predicted by each grid cell, such as 2 in YOLO v1, 5 in YOLO v2, 3 in YOLO v3, C indicates the number of categories of bounding box (no background class) Therefore, for the VOC data set is 20), 5 represents 4 coordinate information and a confidence level.

In the coordinate prediction of bounding box, tx, ty, tw, and th are the predicted outputs of the model, cx and cy represent the coordinates of the grid cell, and pw and ph represent the size of the bounding box before prediction. Bx, by, bw, and bh are the center coordinates and size of the predicted bounding box. The loss of coordinates uses a squared error loss.

\[ b_x = \sigma(t_x) + c_x \]  
\[ b_y = \sigma(t_y) + c_y \]  
\[ b_w = p_w e^{t_w} \]  
\[ b_h = p_h e^{t_h} \]

2.2. Network Structure (Darknet-53)

The network structure Darknet-53 in the YOLOv3 algorithm indicates that there are 53 convolution layers, which actually occupy 74 layers (calculated by the network structure output at run-time). Convolution layer parameter pad in cfg file: If pad is 0, padding is specified by padding parameter; if pad is 1, padding size is size/2 (so that stride==1, the feature image size remains unchanged after
The detection layer is responsible for predicting the box regression value of a certain scale (dividing the number of grids, there are 3 scales 13, 26, 52 respectively predicted at 82, 94, 106 layers) (the regression value of each box predicts 3 boxes includes coordinates, objects and category, a total of $3 \times (4 + 1 + 20) = 75$ values.

![YOLOv3 network structure level diagram](image)

**Figure 1.** YOLOv3 network structure level diagram

On the one hand, the full convolution is basically adopted, and on the other hand, the residual structure is introduced. Darknet-53 is only a feature extraction layer. The source code only uses the convolution layer in front of the pooling layer to extract features. Therefore, multi-scale feature fusion and prediction branches are not reflected in the network structure. The prediction branch adopts a full convolution structure, in which the number of convolution kernels of the last convolution layer is 255, which is 80 classes for the COCO data set: $3 \times (80+4+1)=255$, 3 represents a grid cell Contains 3 bounding boxes, 4 indicates 4 coordinate information of the box, and 1 indicates confidence.
| Type         | Filters Size | Output       |
|--------------|--------------|--------------|
| Convolutional| 32 3˟3       | 256˟256      |
| Convolutional| 64 3˟2/2     | 128˟128      |
| Convolutional| 32 1˟1       |              |
| Convolutional| 64 3˟3       |              |
| Residual     |              | 128˟128      |
| Convolutional| 128 3˟3/2    | 64˟64        |
| Convolutional| 64 1˟1       |              |
| Convolutional| 128 3˟3      |              |
| Residual     |              | 64˟64        |
| Convolutional| 256 3˟3/2    | 32˟32        |
| Convolutional| 128 1˟1      |              |
| Convolutional| 256 3˟3      |              |
| Residual     |              | 32˟32        |
| Convolutional| 512 3˟3/2    | 16˟16        |
| Convolutional| 256 1˟1      |              |
| Convolutional| 512 3˟3      |              |
| Residual     |              | 16˟16        |
| Convolutional| 1024 3˟3/2   | 8˟8          |
| Convolutional| 512 1˟1      |              |
| Convolutional| 1024 3˟3     |              |
| Residual     |              | 8˟8          |
| Avgpool      | Global       |              |
| Connected    | 1000         |              |
| Softmax      |              |              |

**Figure 2.** Darknet-53 model structure

### 2.3. YOLOv3 algorithm network parameter optimization

The YOLOv3 algorithm refers to the anchor parameter, which is a set of a priori boxes with a fixed width and height. In the target detection process, the size of the a priori frame directly affects the speed and accuracy of the detection. In the process of determining the anchor parameter by using K-means and K-means++, two clustering algorithms, in order to reduce the Euclidean distance error caused by the size of the a priori box, the original algorithm is replaced by the cross-match ratio of the label sample box to the a priori box. The Euclidean distance is used as the objective function, and the objective function size indicates the deviation of each sample between the cluster centers. The smaller the objective function value, the better the clustering effect. The calculation formula of the objective function $D$ is:

$$D = \min \sum_{box=0}^{n} \sum_{cen=0}^{k} [1 - IOU_{cen}^{box}]$$ (5)

In the formula, box is the target box of the sample label, cen is the cluster center, n is the number of samples, and k is the number of categories.

### 3. Experiments

#### 3.1. Algorithm implementation environment

This experiment was completed on 1080C with cuda and cudnn accelerated PC. The specific configuration environment is as follows:
3.2. Data set label production

In the target detection problem, the selection of the training data set and the labeling of the original image are two crucial steps. The accuracy of the original image label directly affects the training effect and the accuracy of the test. In the experiment, the data set is made through network search. Firstly, the images in the data set are sorted according to the VOC2007 data set format, and the images in the data set are randomly divided into two categories: training set and test set. Secondly, the labelImg tool is used to mark the images in the training set one by one, and the corresponding target position information file of the xml format is generated. Finally, the python program is written to process the target box position information in xml format and convert it into txt format as the experimental data set label.

3.3. Data set label cluster analysis

Since the VOC data set does not contain target-related data, training with the original parameters of YOLOv3 will have a certain impact on training time and training accuracy. Therefore, the target tag is clustered to obtain a more representative anchor parameter for the target detection.

4. Results and Discuss

The K-means algorithm and the K-means++ algorithm are used to perform dimensional clustering analysis on the target tags. With the difference of k values, the objective function D changes as shown in the following figure:

![Objective function curve](image-url)

**Figure 3. Objective function curve**
It can be seen from the figure that with the increase of k value, the K-means algorithm and K-means++ algorithm objective function values are gradually reduced, and the clustering effect is gradually improved. However, in the process of descending the objective function, the K-means++ algorithm curve is more smoothing, the trend is more stable, and the clustering bias is reduced to some extent, it is better to use the K-means++ algorithm to cluster and obtain the anchor parameter instead of the original Parameters for training and testing.

4.1. In terms of detection speed
Grouping the images in the data set, eliminating some invalid images and using the K-means clustering algorithm to optimize the YOLOv3 network of the anchor parameters (YOLOv3-107) and K-means clustering algorithm to optimize the anchor parameters and improve the same target training set. Network layer number (YOLOv3-101), K-means++ clustering algorithm optimization anchor parameter (YOLOv3-K-107), K-means++ clustering algorithm to optimize anchor parameters and improve network structure (YOLOv3-K-101) four networks Four rounds of experimentation. The experimental results are as follows:

| The Internet       | Average test time | Average miss rate | mAP  |
|--------------------|-------------------|-------------------|------|
| YOLOv3-107         | 0.019s            | 8.6%              | 83.2%|
| YOLOv3-101         | 0.017s            | 0%                | 89.7%|
| YOLOv3-K-107       | 0.019s            | 2.3%              | 91.3%|
| YOLOv3-K-101       | 0.017s            | 0%                | 95.1%|

It can be seen from the table that the network with improved network structure is shortened by 2 ms compared to the unmodified detection average test time, and its other performance is also improved.

4.2. In terms of detection accuracy

![Figure 4. Detection accuracy of different network structures in test set images](image_url)

It can be seen from the figure that the target detection accuracy is higher in different network structures, but the K-means algorithm and the K-means++ algorithm's unimproved network structures YOLOv3-107 and YOLOv3-K-107 all have target detection accuracy. The low case, and the improved
network structure detection has not appeared in this case, indicating that the improved accuracy has improved.

5. Conclusion
Through the results of this experiment, compared with other people's research on image processing in power plants, the YOLOv3 algorithm has greatly improved the speed and accuracy of image processing in power plants. YOLOv3 has an excellent mAP-50 level and is highly capable of positioning bounding boxes. However, in order to ensure real-time operational efficiency, YOLOv3 limits each grid to output only two targets. Looking forward to the future, YOLOv4 and YOLOv5 are constantly advancing and constantly improving to achieve better image processing.

References
[1] Xu Bing. Application of digital image processing in power plant flame detection [D]. North China Electric Power University (Beijing), 2003.
[2] JI Chang-an, ZHANG Xiu-bin, ZHAO Xing-yong, WU Hao, YING Jun-hao, ZENG Guo-hui. Application Research of Image Processing and Fuzzy Control in Power Plant Reconstruction [J]. High Voltage Engineering, 2007 (11): 234 - 239.
[3] Duan Junyong, Wang Shoucheng, Zhao Zenghong, Hao Shiqing. Application of Image Processing Technology in Coal Quality Detection of Power Plants [J]. Journal of Qingdao University of Science and Technology (Natural Science Edition), 2005 (03): 253 - 255.
[4] Appearance inspection of image processing insulators for maintenance inspection of power plants [J]. Electric porcelain arresters, 1994 (05): 29.
[5] LIU Wei, GUO Lijun, ZHANG Rong, HU Yetian. Application of improved YOLOv3 algorithm in container box number positioning [J]. Sensors and Microsystems, 2019, 38 (07): 157 - 160.
[6] Li Jing, Huang Shan. Based on YOLOv3 target tracking method [J/OL]. Electro-optical and control: 1–8 [2019-07-31].
[7] Sheng Yihua, Wu Youxin, Yao Leiyue. A method for detecting illegal parking of shared bicycles based on YOLOv3 [J/OL]. Computer Engineering: 1-8 [2019-07-31].
[8] Ge Wen, Shi Zhengwei. Application of improved YOLOv3 algorithm in pedestrian recognition [J/OL]. Computer Engineering and Applications: 1–11 [2019-07-31].
[9] LIU Xueping, LI Yiqian, LIU Li, WANG Zhe. Improved Yolov3 Target Recognition Algorithm for Adaptive Edge Optimization [J]. Microelectronics & Computer, 2019, 36 (07): 59 - 64.
[10] Li Mao-Hui, Wu Chuan-Ping, Bao Yan, Fang Zhuo-Qun. On the application principle of YOLO algorithm in ma-chine vision. Journal of Computer Applications, 2018, 5 (41): 174 – 176