# Research on Application of Machine Vision in Robots for Live-Power Lines

Wu Han[1], Huang Jianye[1], Lin Shuang[1], Liu Bingqian[1], Fan Yuanliang[1], Chen Haigen[2,*]

[1]Fujian Electric Power Research Institute, FuZhou, 350007, Fujian Province, China
[2]Yangtze Delta Region Institute Of Tsinghua University, Zhejiang Department Of Information Technology, Jiaxing, 314000, Zhejiang Province, China

**Abstract.** The electric fire-exchanging operation is basically used to implement operation such as Wire stripping, drainage wire through the fastening wire clip, clamp bolt fastening, etc. Manual operation is time-consuming and laborious and has very high safety risks, so the automatic electrified ignition robot emerges as The Times require; The automatic operation method should accurately identify the position of wire and drainage line, so as to guide the robot to move and carry out the corresponding operation. Based on this, an intelligent binocular vision guidance method is proposed in this paper. The image data in the working space is collected by the binocular camera in real time, the wire is identified by the image segmentation technology, the wire is separated from the background, and the spatial pose of the wire is solved by the PNP and binocular stereo vision technology. Through YOLOV5 target detection, the position identification and positioning of the thread and clamp of the drainage line are realized, and the automatic fire ignition operation is finally guided by the manipulator. Experiments show that this method has the advantages of high accuracy, good stability and fast speed.

## 1 Introduction

With the increasing improvement of people's living standards, the government bodies, factories, mines and people's lives depend more on electricity. What's more, the contradiction between the need for power outage in power grid transformation and the people's uninterrupted power demand is becoming increasingly prominent, bringing great pressure to the power supply enterprises to fulfill their service commitments. In order to strengthen the daily maintenance, take preventive measures, and achieve timely repair in case of problems, build the power system security defense framework, and deal with various challenges, the power grid company requires that the work on the distribution line must follow the principle of "continuous electricity supply" and be carried out in the form of live line work as far as possible, so as to improve the stability and reliability of power supply.

The common work items of live line work include fire connection, fire cutting, installation of clamp fittings and removal of foreign matters. Generally speaking, these kinds of work account for more than 90% of the daily line live operation. These projects have various operation processes by using different tools. Consequently, a field operator has to accept multi-channel training to master most of the operations, and during his or her operation, mistakes are easy to arise. The live line operation often requires many people to operate on the pole at the same time, such operation is difficult, time-consuming and laborious. Besides, the operators are always in the threat of high voltage and strong electric field, and the frequent takeoff and mental tension are easy to cause personal casualties [1].

In order to meet the needs of economic development, production and people's life for the continuous supply of electricity, the research on live line robot enjoys high application value and prospect. Live line is mainly carried by the mechanical arms with a number of automatic work tools, combined with vision and other multi-sensor fusion technology, to identify, locate and carry out related work on the target object, in which visual assistance serves as the key to help the robot realize automatic fire connection. The vision algorithm proposed in this paper can assist the live line robot, guide the robot to realize such work as automatic wire grasping, wire stripping, wire clip position recognition, robot wire clip fastening. The method in this paper was mainly based on image recognition, detection and segmentation to realize the recognition of wires and clamps. Based on stereo vision and multi view geometry, the pose estimation of objects can be identified. All these provide a strong guarantee for the realization of automatic and intelligent live fire operation.

## 2 Vision system design of robots for live-power lines

The entire robot structure is composed of two

---

* Corresponding author: 35527593@qq.com

manipulators and several working tools. Each manipulator carries a binocular camera. The process of live line operation is as follows:
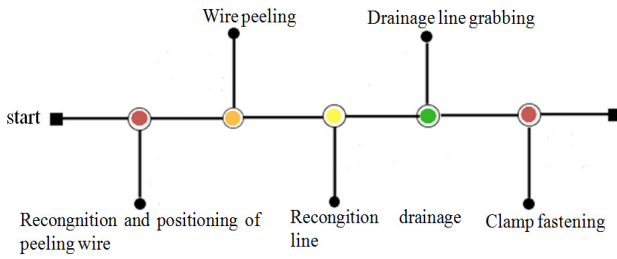


**Figure 1** Process of live line operation

Vision is mainly used to identify and locate the wire to be peeled, identify and locate the leading wire, calculate the spatial coordinates of the wire peeling and find out the position space coordinates, so as to guide the movement and operation of the robot arms.

# 3 Wire identification detection and pose estimation

## 3.1 Segmentation and fitting of peeled wire

Wire segmentation is mainly used for the robot peeling task. The traditional method is usually based on gray image binarization. By setting the threshold and comparing the pixel size, it can be divided into two categories: foreground and background [2,3], but the actual scene is complex, and the threshold is hard to determine. Therefore, this paper uses SOLOV2[4] to segment the target wire, and the network architecture can be seen in Figure 2:
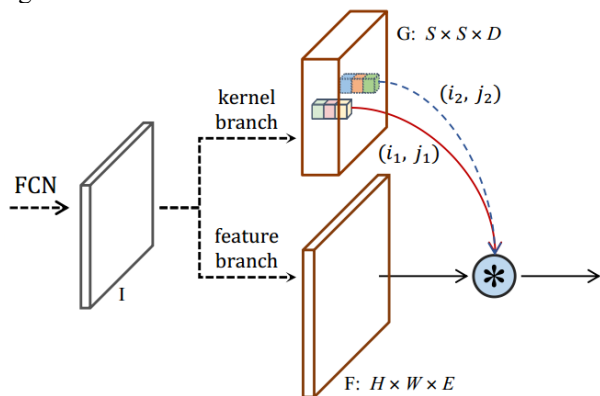


**Figure 2** SOLOV2 network architecture

Among them, $*$ represents dynamic convolution operation. Through dynamic learning, the network divides the mask branch into kernel branch and feature branch to learn convolution kernel and convolution feature respectively. At the same time, non-maximum matrix suppression is used to significantly reduce the reasoning time, so the speed will be improved.

## 3.2 Calculation of pose estimation

According to the rigid body coordinate transformation theory, the camera rotates and translates once according to a certain rotation axis [5,6]

$$X' = RX + T \qquad (1)$$

Where R is the rotation matrix, T the translation vector, and X the coordinate before transformation, and $X'$ the coordinate before transformation.

When the wires are transformed into the wire coordinate system based on their own system, camera A and B are used to calculate the Euler angles of Z-axis and X-axis respectively.

The clockwise rotation is negative, while the counterclockwise one is positive (the rotation direction is the same as that of the right hand system, and it is positive), the results of rotation around non coaxial lines are as follows:

$$R_x = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{bmatrix} \qquad (2)$$

$$R_y = \begin{bmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{bmatrix} \qquad (3)$$

$$R_z = \begin{bmatrix} \cos\gamma & -\sin\gamma & 0 \\ -\sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad (4)$$

When the wire coordinate system is transformed into coordinate system of camera A, $\alpha, \beta, \gamma$ represent the Z-Y-X Euler angle respectively, and the transformation matrix is $R = R_x * R_y * R_z$. If the robot motion transformation matrix R is controlled, the coordinate system of the camera and the peeled wire coordinate system can be ensured to be consistent.

## 3.3 Binocular distance measurement of the wire

In order to cooperate with the robot to finish the wire peeling work and ensure that the cutting point of the robot wire is directly above the robot. The wire pose calculation only needs to change the robot posture, and the translation is completed by the binocular camera to measure the distance of the wire directly above. The binocular ranging idea is mainly based on the triangle similarity principle [7], and uses feature matching to calculate the parallax, as shown in Figure3:
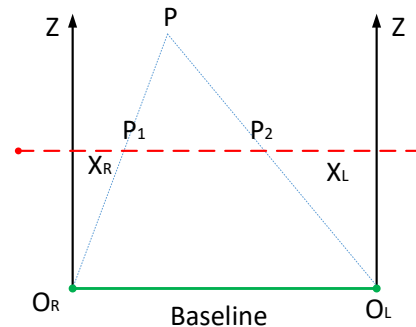


**Figure 3** Stereo System

P is a point in the space, $O_R$ and $O_L$ are the optical

centers of the two cameras, $P_1$ and $P_2$ represent the imaging points of two cameras, f is the focal length, B is the baseline distance of two cameras, according to the trigonometric similarity principle,

$$\frac{B - (X_R - X_L)}{B} = \frac{Z - f}{Z} \qquad (5)$$

we can know:

$$Z = \frac{Bf}{X_R - X_L} \qquad (6)$$

In equation (5), f and B can be obtained by calibration. Therefore, the spatial coordinates (x, y, z) of point P can be acquired only by calculating the offset coordinates (x, y) of parallax A and P pixels relative to the optical center (U, V). Based on the camera calibration focal length [8, 9], and equation (1), the space coordinates (x, y, z) of point P can be calculated.

$$\frac{z}{f} = \frac{X}{x} = \frac{Y}{y} \qquad (7)$$

Based on the pose transformation, the robot can reach the stripping position by moving the distance Z, and rotate the peeling tool to perform the corresponding peeling task.

## 4 Wire clamp threading for visual aid

Wire clamp threading is mainly based on the computer vision system to identify, locate and grasp the target leading wire, and then put it into the fastening clamp with the visual positioning, so as to complete the entire live line work.

### 4.1 Identification and positioning grab of leading wire

The leading wire is a section of arc-shaped conductor, which is often hung on both sides of the tower pole in a hanging shape. And a section of it, stripped and cut, is usually exposed, as shown in Figure 4:
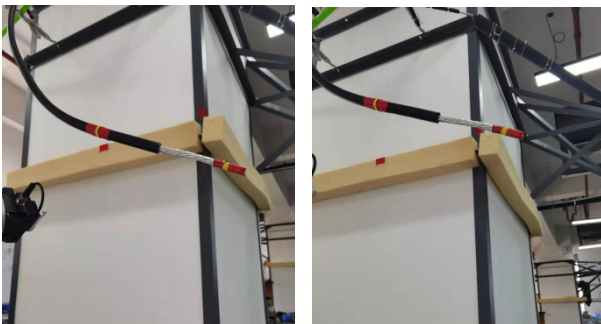


**Figure 4** Drainage line

Leading wire recognition and positioning is mainly used to identify the peeling position and estimate the pose. In this paper, the same as the wire stripping and cutting method, SOLOV2 image segmentation technology is used to train multi classification network and carry on the segment of leading wire.

The leading wire grasping task mainly solves the positioning problem. Generally speaking, the grasping task can be completed when the gripper is under the leading wire and the X or Y axis of the camera (gripper) is parallel to the wire. Based on the example leading wire segmentation, the end position is fitted, and the robot arm is controlled to translate 5cm to the right, down, left and up directions respectively. Besides, the space virtual rectangular frame is drawn, serving as the coordinate plane to establish the user coordinate system. At this time, the pixel coordinates of the wire end in the four images are $P_1$, $P_2$, $P_3$ and $P_4$. PnP is usually used to solve the point motion problem of 3D-2D. That is, the real coordinates of N spatial points and the corresponding pixel coordinates in the known world coordinate system, to calculate the pose of the camera. Solving PnP problems usually includes direct linear transformation, P3P [10], EPNP [11] and UPnP [12]. In addition, there is a least square iterative method based on nonlinear optimization. Zhu Shenglong and other people [13] have used the EPnP algorithm to solve the problem of isolating circuit breaker pose. Similarly, based on the coordinates of four mark points in camera coordinates and the coordinates of virtual coordinate plane mark points, the writer has used EPnP to solve the camera pose.

The marker points used in this calculation are virtual construction and located on the same plane, and the leading wire endpoint is the point coordinate in the world coordinate system, there is no pose, so the manipulator cannot reach the leading wire endpoint through pose transformation. Based on the known length of the exposed part of the leading wire, the endpoint is A, and the exposed wire is B, then $\overrightarrow{BA}$ represents the space vector from B to A. In addition, the position of the camera and the manipulator should be horizontally installed. Consequently, as long as the camera is horizontally installed with the exposed position of the wire through attitude transformation, and the translation transformation solved by EPnP algorithm can reach the grasping position of the leading line. The method is as follows:

1) Calculate the angle $\theta$ between $\overrightarrow{BA}$ and Y axis of the camera in camera coordinate system

$$\theta = \cos^{-1} \frac{\overrightarrow{BA} * \vec{y}}{|\overrightarrow{BA}| * |\overrightarrow{BA}|} \qquad (8)$$

2) Calculate the rotation axis $\vec{n}$ of $\overrightarrow{BA}$ and Y axes. Based on the cross product of $\overrightarrow{BA}$ and $\vec{y}$, let $a = (a_1, a_2, a_3)$, $b = (b_1, b_2, b_3)$, the unit vector of $\overrightarrow{BA}$ and Y axes respectively, then the cross product is expressed as:

$$\vec{n} = \vec{a} \times \vec{b} = \begin{bmatrix} i & j & k \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix} \qquad (9)$$

3) Calculate the rotation vector

$$\vec{R} = \theta\vec{n} \qquad (10)$$

4) Taking the middle position of the leading wire as the grabbing point, the transformation relationship between the camera and the grabbing point is as follows:

$$\begin{bmatrix} X_C \\ Y_C \\ Z_C \\ 1 \end{bmatrix} = \vec{R} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} + \vec{t} \qquad (11)$$

### 4.2 Detection and positioning of clamp target

The target detection mainly detects the target and position of the positioning clamp, which is convenient for the leading wire to pass through. The common target detection algorithms are divided into two-stage detection and one stage detection. The representative algorithms are Fast R-CNN, Faster R-CNN [14,15] and YOLO series [16,17,18,19]. The RCNN series carry the characteristics of high precision and high delay. While the Yolo series algorithms are usually widely used because of the high precision and low delay. Therefore, the writer uses YOLOV5 to realize clamp target detection.

YOLOV5 uses Focus, CBL, CSP1_1 and CSP1_3 as the major network. Among this, CBL, the composite module, is composed of convolution layer, batch regularization and LeakyRelu. Resunit is two CBL composite connection modules, CSP1_X is a composite module composed of CBL, Resunit, convolution layer, batch regularization and leakage correction linear unit. CSP2_X is a composite module of CBL, convolution layer, batch regularization and leakage correction linear unit. Focus is the original image, the slicing process first turns it into 304x304 feature image, and then convolutes it with 32 convolution cores to generate 304x304x32 feature image. This composite module is connected with slices and composed of CbL. SPP (spatial pyramid pooling) is a composite module including CBL and Maxpooling, which is mainly used to solve the defects caused by different image sizes and extract features from various aspects. The detection result of clamp target is shown in Figure 5:
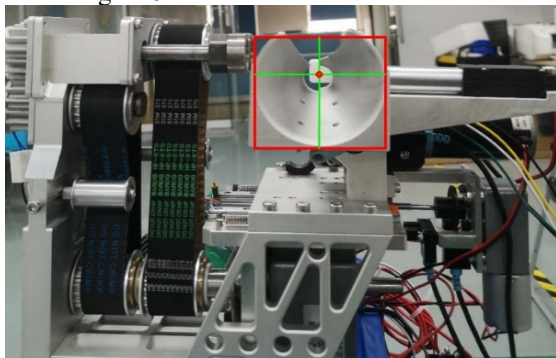


**Figure 5** detection result of clamp target

As shown in the above figure, with the red point in the center of the clamp target as the center, the distance between the upper border and the center point of the target detection is set as $h_1$, the distance between the lower border and the center point $h_2$, the distance between the left border and the center point $w_1$, and the distance between the right border and the center point $w_2$, at this time, the clamp is the threading position in the camera perspective, and the leading wire is pulled into the clamp by the manipulator to complete the threading work.

# 5 Experiment of automatic live line connection under visual guidance

## 5.1 Data collection and processing

In this paper, three kinds of data samples, such as wire to be stripped, drainage wire and clamp, are taken by CCD camera in the experimental site. A large number of image data are collected under different angles, illumination and positions, and 6000 data pictures are processed respectively. As the experimental scene is relatively fixed, in order to make the network model carry better generalization performance, the writer uses such data enhancement methods as random flipping, clipping, HSV color space change and mosaic to amplify the sample data, greatly enriching the sample data. At the same time, labellmg open source data annotation software is used to label the sample data for network training.

## 5.2 Visual guidance for wire peeling

The manipulator A carries a binocular camera, after it collects the image data of the wire to be peeled (as shown in Figure 6), using the image segmentation algorithm SOLOV2 to obtain the example of the wire to be peeled (as shown in Figure 7). Through RANSANC and least square fitting [20] for this example, the straight line of the current wire to be peeled (as shown in Figure 8) can be fitted, and the current angles of 63.21 degrees and 15.02 degrees can be calculated, that is, 63.21 degrees are rotated around the X axis and 15.02 degrees are rotated around the Z axis.
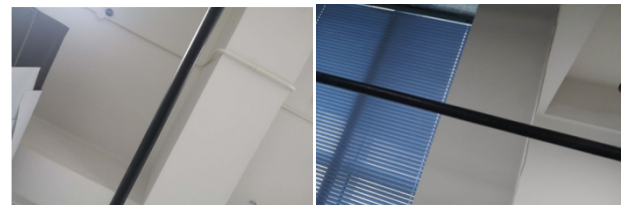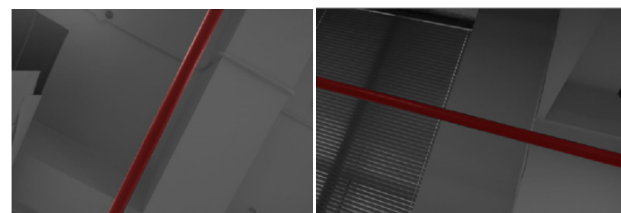


**Figure 6** Peeling wire



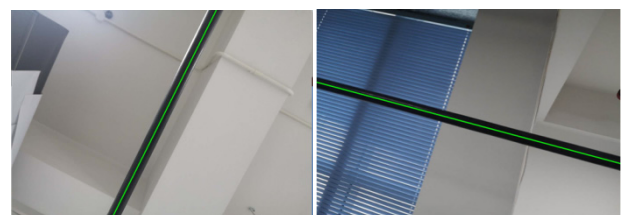**Figure 7** Peeling wire segmentation



**Figure 8** Peeling wire fitting

According to $R = R_x * R_y * R_z$, the rotation matrix R

can be calculated:

$$\begin{bmatrix} 0.455 & -0.827 & 0 \\ -0.862 & 0.429 & 0 \\ 0 & 0 & 0.998 \end{bmatrix}$$

After the motion transformation, adjust manipulator A to change the pose and position. At this time, ensure that the axis Y of the camera plane is parallel to the wire, and axis X is perpendicular to the wire. Meanwhile, according to binocular distance measurement, move along the direction of X, Y and Z axis, so that the wire peeling tool will move to the specified position for wire peeling.

### 5.3 Fixation of lead clamp under visual guidance

The working image data of the current clamp leading line is captured by the camera on manipulator B, and SOLOV2 is used for image segmentation. The results are shown in Figure 9.
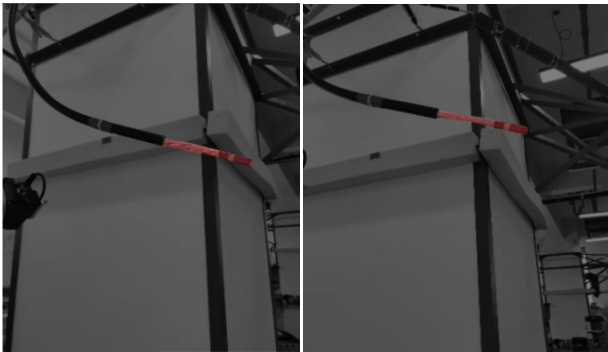


**Figure 9** Result of identification of drainage line

The red area is the segmentation result of lead wire, and the rest is the background. By fitting the segmentation result of lead wire, the PnP pose calculation is carried out to capture the lead wire, and then the clamp target recognition and positioning are performed. In this paper, we use the YOLOV5 algorithm to realize the target detection of the cable clamp. The network training and testing process can be seen in Figure 10. It indicates the relationship between the training set (validation set) loss and the model iteration rounds. As can be seen from the figure, the model starts to converge after about 800 iterations, and performs well in the verification set. The target detection accuracy of the model reaches 99% in the collection data set and 98.5% in the test set, meeting the actual work requirements.
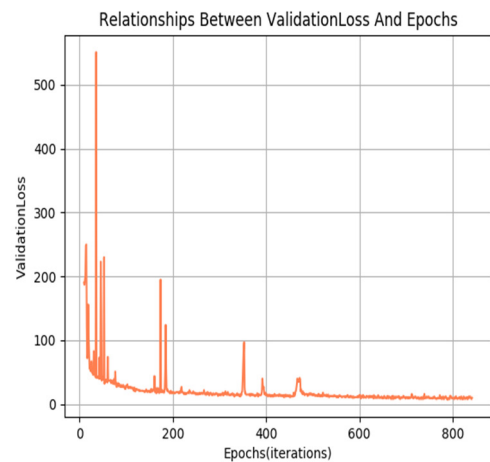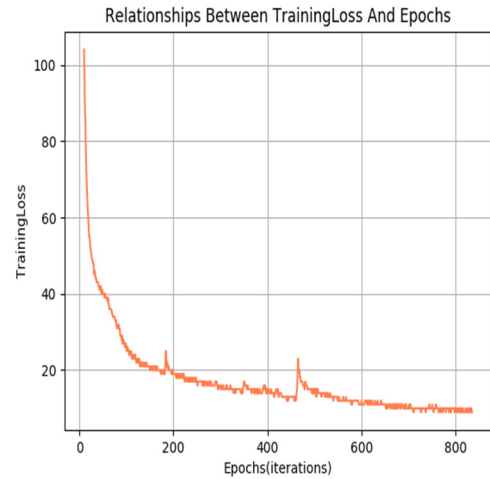


**Figure 10** Model training results (the top one is training results, the bottom is verification results)

## 6 Conclusion

In this paper, according to the actual live line working scene, and the visual guidance of automatic live line operation process, a set of visual intelligent recognition detection methods were designed. Through deep learning, target detection, segmentation, stereo vision and other technologies, target recognition, positioning and other tasks were realized, and the manipulators were guided to complete wire peeling, drainage wire threading, and clamp fastening operations. The experimental results show the effectiveness and reliability of the intelligent visual assistant system designed in this paper, and further demonstrated its practical value in live line operation of distribution network.

## Acknowledgment

## References

1. Shi, L., Bo, L., Gang, H.., et al.(2018)The Hot-line Work Robot for Power Distribution Network. Automation Panorama., 03:68-70.

2. Badekas, E., Papamarkos, N. (2009) Estimation of appropriate parameter values for document binarization techniques. International Journal of Robotics and Automation.,24(1): 66-78.

3. Sauvola, J., Pietikainen, M. (2000) Adaptive document image binarization. Pattern Recognition., 33(22): 225-236.

4. Wang, X., R, Zhang., Kong, T., et al. SOLOv2: (2020) Dynamic, Faster and Stronger.

5. K Tchoń. (2000). Singularities of the euler wrist. Mechanism & Machine Theory., 35(4), 505-515.

6. Fan, Z., XiBin, C. (2002) A New Large-scale Transformation Algorithm of Quaternion to Euler Angle. Journal of Nanjing University of Science And Technology., (04).

7. XueJun, Q., Lu, Z. (2011)3D Measurement Method Based on Binocular Vision Technique[J]. Computer Simulation., 28(02):373-377.

8. Liang, R., TianShuang, Q. (2014) Camera Calibration Method Based on Three Orthogonal Constraints, Acta Electronica Sinica., 42(10): 2050-2054.

9. ChengZhi, S., GuoEn, W., JiangTao, H., et al. (2011) Distortion correction for images in planar metrology. Optics and Precision Engineering., 19(01):161-167.

10. X.-S. Gao, X.-R. Hou, J. Tang, and H.-F. Cheng, (2003) Complete solution classifcation for the perspective-three-point problem, IEEE Transactions on Pattern Analysis and Machine Intelligence, 25:930–943

11. Lepetit, V., Moreno-Noguer, F., Fua, P. (2009) Epnp: an accurate o(n) solution to the pnp problem. International Journal of Computer Vision, 81(2), 155-166.

12. Penate-Sanchez, A., Andrade-Cetto, J., Moreno-Noguer, F. (2013). Exhaustive linearization for robust camera pose and focal length estimation. IEEE Transactions on Pattern Analysis & Machine Intelligence., 35(10), 2387-2400.

13. ShengLong, Z., BinBin, L., ChenChen, Z., et al. (2020) Study on visual guidance and positioning method of automatic make-up and disassembly robot of isolated circuit breaker., 56(07):205-211.

14. Girshick, R. (2015) Fast r-cnn. Computer Science.

15. Ren, S., He, K., Girshick, R., Sun, J. (2017) Faster r-cnn: towards real-time object detection with region proposal networks. IEEE Transactions on Pattern Analysis & Machine Intelligence., 39(6), 1137-1149.

16. Redmon, J., Divvala, S., Girshick, R., et,al. (2016)You Only Look Once: Unified, Real-Time Object Detection. Computer Vision & Pattern Recognition. IEEE.

17. Redmon, J., Farhadi, A. (2017) YOLO9000: Better, Faster, Stronger. IEEE Conference on Computer Vision & Pattern Recognition (pp.6517-6525). IEEE.

18. Redmon J, Farhadi A. (2018) YOLOv3: An Incremental Improvement. https://arxiv.org/abs/1804.02767.

19. Bochkovskiy, A., Wang, C. Y., Liao, H,. (2020) Yolov4: optimal speed and accuracy of object detection.

20. YiBin, Y., ShuHua, H., Jian, K., et al. (2014) Total Least Squares Algorithm for Fitting Spatial Straight Lines. Geomatics and Information Science of Wuhan University.,39(05):571-574.