Binocular vision mechanical arm system based on salient region target recognition

Ronghua Li\textsuperscript{1a}, Ruiliang Wang\textsuperscript{2a}, Xueying Tian\textsuperscript{3a}
\textsuperscript{a}Mechanical Engineering, Dalian Jiaotong University, Dalian, Liaoning province, 116028, China
\textsuperscript{1}lironghua705@163.com, \textsuperscript{2}545306723@qq.com, \textsuperscript{3}1175253115@qq.com

Abstract: Robotic arms have been widely used in the sorting of manufacturing lines. During the sorting process of the robot arm, the captured speed and accuracy are core problems. In order to improve that, this paper proposes a binocular vision sorting robot arm system based on the target identification of significant regions. First, establish the target artifact training set and the Faster R-CNN algorithm is imported. A binocular camera is used to take pictures of the workpiece in the tray, and then deep learning algorithm is used to extract the significant area where the object is located. Image processing is performed on this region. Then, the three-dimensional coordinates of the workpiece centroid are calculated and passed to the manipulator for grasping and sorting. Because the significance area is extracted, the amount of calculation by the image processing algorithm is greatly reduced, and the speed is significantly improved. Finally, the picking and sorting experiment was designed. The results show that the system has high accuracy, strong robustness and strong portability.

1. Introduction
With the rapid development of the manufacturing industry, industrial robots have been widely used. Among them, the sorting operation is the basic link in the production line. In the sorting operation of the robot arm, because the base speed of the conveyor belt is high, the manipulator must be able to quickly identify and locate the target to grasp. As a result, many scholars have carried out a lot of work and research, such as Wang Peng et al\textsuperscript{[1]} developed a sort of coal gangue based on rock texture features, using image processing and pattern recognition technology; Zhao Lin et al\textsuperscript{[2]} using Open MV machine vision method to realized the recognition of objects and captures them. However, the existing sorting robot arm works with techniques such as color sensor, two-dimensional code recognition, and template matching, which still has some problems such as low sorting efficiency, large waste of resources, and poor robustness.

This paper proposes a robotic arm control system that combines deep learning with image processing algorithms. To begin with, extract the significance area where the target work piece is located, and then the image of the area is processed. The aim is to combine the two advantages which are the fast target identification of deep learning and fast resolution of the target object’s centroid coordinates of the image processing algorithm, and then build a faster and more accurate robotic arm sorting system.
2. The Significance region extraction algorithm based on Deep Learning

2.1 Make training data sets
In order to enable the mechanical arm sorting system to extract the significance region where the target work piece is located, the training set is firstly established for the deep learning algorithm. The target work piece used in this paper is the brake slide valve with complex contour, and the interfering object is the common building material accessories of similar materials, as shown in figure 1.

VOC2007 was selected in this paper to produce the data set. The source images were manually shot in the mode of continuous shooting by camera, with a total of 240 images, including 60 training sets, 120 test sets and 60 verification sets. After selecting the training set photos, manual marquee labeling is started with the LabelImg tool. The process of source photo set and manual labeling is shown in figure 2. The above annotation data is packaged in Python code to compete the generation of the training dataset.

2.2 Significance area extraction
The above training set data were transferred to the Faster R-CNN network for training, in which the model defined by the Faster RCNN VGG16 interface, the constants used to define the training and the optimization algorithm were used, and weight attenuation operation was added. The training USES the multi-process iterator interface, which makes more use of the resources of the computer, and then USES the updater to associate the training data with the optimization algorithm. The final detection effect of the image taken by binocular camera is shown in figure 3. The visible target work piece is effectively identified, which creates favorable conditions for the speed up of subsequent image processing.

3. Significance region image processing algorithm
In this paper, through deep learning algorithm, the significance area where the target work piece is located is extracted in advance and marked with a rectangular box. Then, only image processing of this area can significantly reduce the calculation time and improve the sorting speed of the mechanical arm.

3.1 Image preprocessing
In order to find the centrograph of the target workpiece, series of image processing algorithms were taken to extract its features, including mask, threshold segmentation, morphological transformation, boundary extraction, Hough corner point detection and line detection. The processing effect is shown in figure 4.
3.2 Centroid calculation

It can be seen that since this workpiece is a complex polygon, after feature extraction, no simple graph is formed to directly obtain centroid. To extract the centroid of those complex polygon, the method of maximum inner is effective. The center of the rectangular box above is used as the starting point for center diffusion. The resulting inner rectangle then traverses the 8 neighborhood at the center of the smallest outer rectangle. Take the rectangle with the largest area as the largest inner rectangle, and draw its centroid. As shown in figure 5.

4. Mechanical Arm Sorting System

4.1 Mechanical Arm System

The mechanical arm used in this system is Dobot mechanical arm, which is a tabletop type four-axis mechanical arm, with a repetition accuracy of 0.2mm. The manipulator has four degrees of freedom to enabling the manipulator to achieve rapid and accurate positioning in horizontal and vertical planes. A suction cup can be installed at the end of the robot arm to control the suction and release action through the switch volume, realizing the sorting work.

The camera used in this system is the fixed baseline USB3.0 binocular vision. The resolution of the left and right lenses of the camera is 1280*720 respectively, and each camera has 100,000 pixels. The exposure can be automatically adjusted within a certain range according to the current ambient brightness. It can meet the requirements of the experimental system. Both of the mechanical arm and camera are shown as figure 6.

4.2 Experimental System Construction

The mechanical arm and the binocular camera are connected to the computer and controlled by unified Python code. The establishment of experimental platform is shown in figure 7.
Figure 7 Schematic diagram of the experimental system

The experimental process is as follow: the workpiece is randomly placed in the working area. After the mechanical arm is initialized, the workpiece is photographed by binocular camera. Then the computer can identify the target workpiece and extract its centroid coordinates by the algorithm in this paper. Then the coordinate value is converted to the manipulator coordinate system, and the manipulator grabs it.

4.3 Analysis of Experimental Results

In 50 grasping experiments, the target workpiece was accurately identified. But there is some error between the calculated coordinates and the real value. In this paper, the errors of X-axis, Y-axis and Z-axis were calculated, and the broken line diagram was made as shown in figure 8. The error of Z-axis is relatively small, while that of X-axis and Y-axis is relatively large. From the overall grasping effect, the accuracy is relatively high, the system has a strong robustness.

Figure 8 Capture error statistics of mechanical arm

5. Conclusion

Through experimental verification, the binocular visual mechanical arm sorting system designed in this paper based on deep learning basically achieves the expected effect. Moreover, if the initial training data set is changed into other artifacts, it can be transplanted and used for grasping other targets. In the following research, it is planned to add a conveyor belt kit, place the work piece on the conveyor belt, and truly simulate the grasping conditions of the real production line, and continue to further study on improving the algorithm speed and accuracy of calculating the target centroid coordinates.

Acknowledgments

This work has been supported by National key R & D plan(2018YFB1107805); the Liaoning Provincial Natural Science Foundation Project (20180520015); and the Dalian Support for High-Level Talent Innovation and Entrepreneurship Project (2017RQ134)

Reference:

[1] Wang P;Cao X;Xia J;WU XD;Research on multi-manipulator coal and gangue sorting robot system based on machine vision[J],Industry and Mine Automation,2019.9,1-7

[2] Zhao L;Gao Z;Yang X;Liu J;Xue J.Target grasping and sorting based on Open MV vision[D],The 16th shenyang scientific academic annual conference,2019
[3] Zhou M; Zhang D; Zhou T; Chen X. Control of picking robot manipulator based on binocular vision [J]. Computer & Digital Engineering, 2018-11

[4] Chen Y; Bian X. Medical waste classification system based on machine vision and deep learning [J]. Computer Programming Skills & Maintenance, 2019, No.407, 110-112

[5] Bai C; Yan Z; Song J. Visual grasp control of robotic arm based on deep learning [J]. Manned Spaceflight, 2018.6

[6] Chen L; Nie X. Robot recognition, location and grabbing system based on binocular vision [J]. Journal of Mechanical & Electrical Engineering, 2019.08-11

[7] Guo Y, Zhao D. Optimization design of non-coplanar target for pose measurement with monocular vision system [J]. Optik - International Journal for Light and Electron Optics, 2017, 131: 72-78.

[8] Pérez L, Rodríguez N, et al. Robot Guidance Using Machine Vision Techniques in Industrial Environments: A Comparative Review [J]. Sensors, 2016, 16(3): 335.